

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 04:00:18 ; Search time 4859 Seconds
(without alignments)
10552.544 Million cell updates/sec

Title: US-09-939-853A-74
Perfect score: 1183
Sequence: 1 agctagagctccaggacc.....tctcttgatgatgctag 1183

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_scs.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1183	100.0	1183	6	AX443133	Sequence
2	1183	100.0	1183	6	AX443135	Sequence
3	1132	95.7	2567	6	AX452880	Sequence
4	1046	88.4	2538	9	BC042041	Homo sapi
5	735	62.1	786	6	AX511153	Sequence
6	735	62.1	786	6	AX572845	Sequence
7	735	62.1	786	9	AF290985	Homo sapi
8	735	62.1	786	9	AF326353	Homo sapi
9	735	62.1	2415	9	AK025645	Homo sapi
10	684	57.8	2788	6	AX780857	Sequence
11	483	40.8	737	6	AX511155	Sequence
12	483	40.8	737	9	AF290986	Homo sapi
13	354	29.9	66741	9	HS460J8	Human DNA
14	354	29.9	145833	2	AC026539	Homo sapi
15	152	12.8	145068	9	HS0977B1	Human DNA
16	152	12.8	145833	2	AC026539	Homo sapi
17	43	3.6	777	6	AX511151	Sequence
18	43	3.6	1321	10	AF434990	Mus muscu
19	43	3.6	1348	6	AX511150	Sequence
20	43	3.6	1348	10	AF287467	Mus muscu
21	43	3.6	1384	10	BC052655	Mus muscu
22	41	3.5	163240	10	AL935150	Mouse DNA
23	30	2.5	30	6	AX452884	Sequence
24	29	2.5	299947	2	AC125701	Rattus no
25	29	2.5	322940	2	AC123560	Rattus no
26	26	2.2	26	6	AX443200	Sequence
27	25	2.1	25	6	AX511164	Sequence
28	25	2.1	25	6	AX511165	Sequence
29	25	2.1	25	6	AX511166	Sequence
30	23	1.9	23	6	AX511163	Sequence
31	23	1.9	8028	2	AC014511	Drosophila
32	23	1.9	14560	10	AB011527	Rattus no
33	23	1.9	177577	3	AC022349	Drosophila
34	23	1.9	179016	3	AC018489	Drosophila
35	23	1.9	189516	2	AC128065	Rattus no
36	23	1.9	270729	2	AC133702	Rattus no
37	23	1.9	300469	3	AE003503	Drosophila
38	22	1.9	22	6	AX443201	Sequence
39	22	1.9	475	9	HSPPROT04	Human P pro
40	22	1.9	3070	6	E24027	Method for
41	22	1.9	3070	6	E24028	Method for
42	22	1.9	3070	9	HUMPPRO	Homo sapien
43	22	1.9	3082	9	BC012097	Homo sapi
44	22	1.9	67718	2	AC135349	Homo sapi
45	22	1.9	67718	2	AC135349	Homo sapi
46	22	1.9	155749	2	AC135345	Homo sapi
47	22	1.9	160842	2	AC124090	Homo sapi
48	22	1.9	177613	2	AC124089	Homo sapi
49	22	1.9	180546	2	AC017046	Homo sapi
50	22	1.9	188766	9	AC135329	Homo sapi
51	22	1.9	192087	9	AC121324	Homo sapi
52	22	1.9	195397	2	AC135718	Homo sapi
53	22	1.9	195895	2	AC117033	Rattus no
54	22	1.9	196827	9	AC124091	Homo sapi
55	22	1.9	215835	2	AC098127	Rattus no
56	22	1.9	224266	2	AC127197	Rattus no
57	22	1.9	238907	2	AC128162	Rattus no
58	21	1.8	21	6	AX452882	Sequence
59	21	1.8	110000	2	AC103328	Continuation (2 of
60	21	1.8	151416	2	AC137060	Bos tauru
61	21	1.8	156129	2	AC119240	Mus muscu
62	21	1.8	169618	10	AC122811	Mus muscu
63	21	1.8	178293	2	AC144901	Sus scrofa
64	21	1.8	181245	2	AC137532	Bos tauru
65	21	1.8	186923	2	AC140241	Mus muscu

66	21	1.8	188843	2	AC139844	Mus muscu	20	C 139	1.7	156563	2	AC079458	Homo sapi	
67	21	1.8	198893	2	AC133116	Rattus no	20	C 140	20	1.7	156578	9	AC102802	Homo sapi
68	21	1.8	206826	2	AC145413	Sus scrofa	20	C 141	20	1.7	157120	9	CNS01DXB	Human chr
69	21	1.8	213182	2	AC098767	Rattus no	20	C 142	20	1.7	157289	2	AC021447	Homo sapi
70	21	1.8	222922	2	AC118839	Rattus no	20	C 143	20	1.7	157579	9	AC018639	Human chr
71	21	1.8	224989	2	AC133973	Rattus no	20	C 144	20	1.7	160690	2	AC023457	Homo sapi
72	21	1.8	249671	2	AC130977	Rattus no	20	C 145	20	1.7	162342	2	AP004687	Oryza sat
73	21	1.8	254581	2	AC102993	Rattus no	20	C 146	20	1.7	162564	10	AC127575	Mus muscu
74	21	1.8	261476	2	AC129130	Rattus no	20	C 147	20	1.7	162808	2	AC021340	Homo sapi
75	21	1.8	270006	2	AC129130	Rattus no	20	C 148	20	1.7	163132	9	AC025038	Homo sapi
76	21	1.8	272636	2	AC108647	Rattus no	20	C 149	20	1.7	167815	5	AC114888	Pan trogl
77	21	1.8	337732	2	AC105843	Rattus no	20	C 150	20	1.7	169095	5	EX088685	Febratfish
78	20	1.7	20	6	AX443199	Sequence	20	C 151	20	1.7	172749	9	AC009731	Homo sapi
79	20	1.7	20	6	AX452883	Sequence	20	C 152	20	1.7	173591	2	AC140351	Mus muscu
80	20	1.7	277	6	AX090683	Sequence	20	C 153	20	1.7	175840	9	AC019306	Homo sapi
81	20	1.7	277	6	BD045216	Sequence	20	C 154	20	1.7	178255	9	AC019197	Homo sapi
82	20	1.7	525	6	AX867799	Sequence	20	C 155	20	1.7	180631	9	AC107482	Homo sapi
83	20	1.7	525	6	BD147861	Primer fo	20	C 156	20	1.7	181320	2	AC092060	Homo sapi
84	20	1.7	539	6	AR418884	Sequence	20	C 157	20	1.7	184516	2	AC009629	Homo sapi
85	20	1.7	539	6	BD114437	Sequence	20	C 158	20	1.7	184541	2	AC011405	Homo sapi
86	20	1.7	675	6	AR016398	Sequence	20	C 159	20	1.7	184824	2	AC073738	Homo sapi
87	20	1.7	675	6	AR019256	Sequence	20	C 160	20	1.7	185332	9	AC016673	Homo sapi
88	20	1.7	1047	6	AX653434	Sequence	20	C 161	20	1.7	187752	10	AC124689	Mus muscu
89	20	1.7	1491	9	AR228313	Homo sapi	20	C 162	20	1.7	189707	10	AC130551	Mus muscu
90	20	1.7	1527	9	SSC277921	Samirir s	20	C 163	20	1.7	191950	2	AC022307	Homo sapi
91	20	1.7	1530	6	AX695859	Sequence	20	C 164	20	1.7	194904	2	AC128309	Rattus no
92	20	1.7	1530	6	AX695859	Sequence	20	C 165	20	1.7	196953	9	AC009623	Homo sapi
93	20	1.7	1533	6	HSP320182	Sequence	20	C 166	20	1.7	197006	2	AC102867	Mus muscu
94	20	1.7	1589	9	HSU07236	Human mutan	20	C 167	20	1.7	199014	2	AC073849	Homo sapi
95	20	1.7	1620	12	AY335586	Synthetic	20	C 168	20	1.7	200697	2	AC119512	Rattus no
96	20	1.7	1707	3	AX114952	Ciona int	20	C 169	20	1.7	201327	2	AC108443	Mus muscu
97	20	1.7	2032	6	AX695858	Sequence	20	C 170	20	1.7	202827	9	AC016907	Homo sapi
98	20	1.7	2032	6	BD223388	Method of	20	C 171	20	1.7	202827	9	AC016907	Homo sapi
99	20	1.7	2032	6	HS1CKE	Human lck m	20	C 172	20	1.7	203257	10	AL671571	Mouse DNA
100	20	1.7	2032	9	HMLCKAA	Human lymph	20	C 173	20	1.7	206383	10	AC008160	Mus muscu
101	20	1.7	2041	9	HSTCTPK	X50527 Human T cel	20	C 174	20	1.7	215875	2	AC104611	Rattus no
102	20	1.7	2129	6	AR270889	Sequence	20	C 175	20	1.7	217141	2	AC073998	Homo sapi
103	20	1.7	2169	10	RA1P13KB	Human T-lym	20	C 176	20	1.7	217615	9	AC011005	Homo sapi
104	20	1.7	2182	9	BC013200	BC013200 Homo sapi	20	C 177	20	1.7	217685	2	AC123225	Rattus no
105	20	1.7	2915	9	AX055368	Human lymph	20	C 178	20	1.7	231741	2	AC110855	Rattus no
106	20	1.7	3075	9	HS1CK1	X14053 Human lymph	20	C 179	20	1.7	233036	2	AC146694	Rattus no
107	20	1.7	3157	10	NMP3KB	Y13569 Mus musculus	20	C 180	20	1.7	235436	2	AC128048	Rattus no
108	20	1.7	3194	10	BC006796	BC006796 Mus muscu	20	C 181	20	1.7	238887	2	AC109745	Rattus no
109	20	1.7	3270	8	AY035059	AY035059 Arabidops	20	C 182	20	1.7	238965	2	AC126647	Rattus no
110	20	1.7	3701	5	CHTKCKL	J03579 Chicken tyr	20	C 183	20	1.7	239583	2	AC124977	Mus muscu
111	20	1.7	4073	5	GG1K1	X60380 G.gallus tk	20	C 184	20	1.7	240364	2	AC114875	Rattus no
112	20	1.7	31842	6	AX695857	Sequence	20	C 185	20	1.7	241288	2	AC133712	Rattus no
113	20	1.7	35730	9	HGN74G7	Z69715 Human DNA s	20	C 186	20	1.7	241288	2	AC133712	Rattus no
114	20	1.7	59634	9	AC062028	AC062028 Homo sapi	20	C 187	20	1.7	243508	2	AC137584	Mus muscu
115	20	1.7	61515	9	HSBJ675E8	AL121991 Human DNA	20	C 188	20	1.7	246924	2	AC120914	Rattus no
116	20	1.7	63540	2	AC101468	AC101468 Mus muscu	20	C 189	20	1.7	249624	2	AC109677	Rattus no
117	20	1.7	64016	2	AC101180	AC101180 Homo sapi	20	C 190	20	1.7	258888	2	AC109677	Rattus no
118	20	1.7	64545	2	AC101180	AC101180 Homo sapi	20	C 191	20	1.7	263984	2	AC132831	Rattus no
119	20	1.7	72393	2	AC101182	AC101182 Mus muscu	20	C 192	20	1.7	264928	10	AC131028	Mus muscu
120	20	1.7	72393	2	AC101182	AC101182 Mus muscu	20	C 193	20	1.7	269576	2	AC073684	Mus muscu
121	20	1.7	76332	2	AC021312	AC021312 Homo sapi	20	C 194	20	1.7	271932	2	AC109948	Rattus no
122	20	1.7	84411	8	AT78H10	AL133248 Arabidops	20	C 195	20	1.7	286655	2	AC098489	Rattus no
123	20	1.7	87503	8	AT78H10	AL133248 Arabidops	20	C 196	20	1.7	349970	1	EX571659	Wolinnella
124	20	1.7	109359	9	AR406002_4	Continuation (5 of	20	C 197	20	1.7	349980	6	AX656302	Sequence
125	20	1.7	110000	6	AR406002_4	Continuation (5 of	20	C 198	20	1.6	262	11	G11894	human STS W
126	20	1.7	114800	9	AC123367	AC123367 Homo sapi	20	C 199	20	1.6	688	9	HSJ329069	Homo sapi
127	20	1.7	120194	9	AL355474	AL355474 Human DNA	20	C 200	20	1.6	732	10	AF507943	Rattus no
128	20	1.7	120739	2	AL355474	AL355474 Human DNA	20	C 201	20	1.6	754	9	H006641S18	Rattus no
129	20	1.7	145476	9	HS246H3	AF305874 Homo sapi	20	C 202	20	1.6	758	9	H006641S18	Rattus no
130	20	1.7	145476	9	HS246H3	AF305874 Homo sapi	20	C 203	20	1.6	853	11	BV054821	Homo sapi
131	20	1.7	148266	2	AC087532	AC087532 Homo sapi	20	C 204	20	1.6	1143	9	AB081120	Homo sapi
132	20	1.7	149281	2	AC090903	AC090903 Homo sapi	20	C 205	20	1.6	1149	9	AY237799	Homo sapi
133	20	1.7	151628	2	AC016665	AC016665 Homo sapi	20	C 206	20	1.6	1467	6	AR268796	Sequence
134	20	1.7	151628	2	AC016665	AC016665 Homo sapi	20	C 207	20	1.6	1548	6	BD191521	Diagnosis
135	20	1.7	151874	2	AC011268	AC011268 Homo sapi	20	C 208	20	1.6	1556	10	AF146738	Rattus no
136	20	1.7	152054	10	AC122009	AC122009 Mus muscu	20	C 209	20	1.6	1580	6	AX527771	Sequence
137	20	1.7	156183	10	AL663051	AL663051 Mouse DNA	20	C 210	20	1.6	1718	10	AF007016	Vibrio ch
138	20	1.7	156280	9	AC079595	AC079595 Homo sapi	20	C 211	20	1.6	1727	10	AF075266	Mus muscu
139	20	1.7	156280	9	AC079595	AC079595 Homo sapi	20	C 211	20	1.6	1761	6	AX577934	Sequence

212	19	1.6	1791	8	LEU18678	UI8678	Lycopersico
213	19	1.6	1810	3	AF519443	AF519443	Percolomo
214	19	1.6	1833	6	AX622893	AX622893	Sequence
215	19	1.6	1841	10	AF075265	AF075265	Mus muscu
216	19	1.6	1865	5	BC043740	BC043740	Xenopus l
217	19	1.6	1890	10	AF075264	AF075264	Mus muscu
218	19	1.6	1904	10	BC010748	BC010748	Mus muscu
219	19	1.6	1943	10	AF075267	AF075267	Mus muscu
220	19	1.6	1950	5	BC045462	BC045462	Danio rer
221	19	1.6	1974	10	AF075263	AF075263	Mus muscu
222	19	1.6	1982	5	GGYRA	GGYRA	G.gallus yr
223	19	1.6	2004	10	AF075262	AF075262	Mus muscu
224	19	1.6	2120	6	AX713893	AX713893	Sequence
225	19	1.6	2120	6	AK055798	AK055798	Homo sapi
226	19	1.6	2124	4	BTWRLLG	BTWRLLG	Bovine mRNA
227	19	1.6	2142	9	BSA2CHIA	BSA2CHIA	H.sapiens a
228	19	1.6	2155	9	BC056757	BC056757	Homo sapi
229	19	1.6	2172	10	AK128907	AK128907	Mus muscu
230	19	1.6	2197	4	BTTRS2	BTTRS2	Bovine mRNA
231	19	1.6	2315	9	BC011393	BC011393	Homo sapi
232	19	1.6	2335	10	RNU12973	RNU12973	Rattus norv
233	19	1.6	2393	10	MUSTSPK	MUSTSPK	Mouse mRNA
234	19	1.6	2440	10	MSRM03591	MSRM03591	Mouse mRNA
235	19	1.6	2560	10	MUSSRM	MUSSRM	Mouse mRNA
236	19	1.6	2560	10	BC029546	BC029546	Mus muscu
237	19	1.6	2597	10	MUSPKRBP	MUSPKRBP	Mus muscu
238	19	1.6	2771	6	AR270538	AR270538	Sequence
239	19	1.6	2771	6	HSU43408	HSU43408	Human tyros
240	19	1.6	2792	9	BC035782	BC035782	Homo sapi
241	19	1.6	3146	10	BC008617	BC008617	Mus muscu
242	19	1.6	3670	10	BC063081	BC063081	Mus muscu
243	19	1.6	3670	10	BC043087	BC043087	Mus muscu
244	19	1.6	3681	10	AF295638	AF295638	Mus muscu
245	19	1.6	3880	8	AK121394	AK121394	Oryza sat
246	19	1.6	4410	10	AF456412	AF456412	Mus muscu
247	19	1.6	4871	10	RNU28938	RNU28938	Rattus norv
248	19	1.6	5075	10	BC052743	BC052743	Mus muscu
249	19	1.6	5106	10	AF135166	AF135166	Mus muscu
250	19	1.6	5907	10	RATCP801	RATCP801	Rattus norv
251	19	1.6	9045	10	MAU71280	MAU71280	Mesocricetu
252	19	1.6	10511	9	AF097738	AF097738	Homo sapi
253	19	1.6	10759	1	AB004143	AB004143	Vibrio ch
254	19	1.6	19677	1	AF196903	AF196903	Homo sapi
255	19	1.6	26453	8	SPCC895	SPCC895	S.pombe c
256	19	1.6	33414	9	HSU65590	HSU65590	Homo sapien
257	19	1.6	37752	9	HS361H4C	HS361H4C	Human DNA
258	19	1.6	43878	9	BS284655	BS284655	Human DNA
259	19	1.6	50087	2	AC132823	AC132823	Homo sapi
260	19	1.6	55438	9	AC139752	AC139752	Homo sapi
261	19	1.6	58001	9	AC092978	AC092978	Homo sapi
262	19	1.6	59257	9	AC106019	AC106019	Homo sapi
263	19	1.6	62456	2	AC099992	AC099992	Mus muscu
264	19	1.6	69216	9	AL590230	AL590230	Human DNA
265	19	1.6	69345	2	AC105016	AC105016	Homo sapi
266	19	1.6	70329	2	AC087718	AC087718	Homo sapi
267	19	1.6	70869	9	AC024704	AC024704	Homo sapi
268	19	1.6	73503	10	AL645603	AL645603	Mouse DNA
269	19	1.6	74968	8	AB019230	AB019230	Arabidops
270	19	1.6	75877	9	AP005401	AP005401	Homo sapi
271	19	1.6	76640	9	BS511022	BS511022	Human DNA
272	19	1.6	80141	9	HSEW8GAR	HSEW8GAR	Homo sapien
273	19	1.6	84492	2	AC021676	AC021676	Homo sapi
274	19	1.6	87857	2	AC139662	AC139662	Homo sapi
275	19	1.6	90907	9	AL513543	AL513543	Human DNA
276	19	1.6	93100	8	AP003914	AP003914	Oryza sat
277	19	1.6	102530	2	AC147060	AC147060	Homo sapi
278	19	1.6	106172	9	AL591122	AL591122	Human DNA
279	19	1.6	109296	9	AC083949	AC083949	Homo sapi
280	19	1.6	109891	9	AL353897	AL353897	Human DNA
281	19	1.6	110000	2	AL139250	AL139250	Homo sapi
282	19	1.6	110144	9	AC008053	AC008053	Homo sapi
283	19	1.6	110309	10	AL592404	AL592404	Mouse DNA
284	19	1.6	110309	10	AL592404	AL592404	Mouse DNA
285	19	1.6	111746	9	HS327016	HS327016	Human DNA
286	19	1.6	113196	9	HSJ697K14	HSJ697K14	Human DNA
287	19	1.6	115746	9	BX248579	BX248579	Human DNA
288	19	1.6	117071	9	AL133375	AL133375	Human DNA
289	19	1.6	120027	2	AC020981	AC020981	Homo sapi
290	19	1.6	121799	2	AC105930	AC105930	Magnaport
291	19	1.6	126312	2	AC000026	AC000026	Homo sapi
292	19	1.6	129414	2	AC068398	AC068398	Homo sapi
293	19	1.6	132023	9	AC020596	AC020596	Homo sapi
294	19	1.6	133095	2	AC046146	AC046146	Mus muscu
295	19	1.6	134019	2	AC021361	AC021361	Homo sapi
296	19	1.6	136511	2	AC145621	AC145621	Homo sapi
297	19	1.6	137259	2	AC010392	AC010392	Homo sapi
298	19	1.6	137718	9	AL137218	AL137218	Human DNA
299	19	1.6	138483	2	AC135556	AC135556	Gallus ga
300	19	1.6	139625	2	AC141120	AC141120	Rattus no
ALIGNMENTS							
RESULT 1							
AX443133							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
Homo sapiens (human)							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE							
AUTHORS							
Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R.,							
Shinkens, R.A., Sytek, K.A., Szekeres, E.S., Tomlinson, J.E.,							
Topper, J.N. and Yang, R.B.							
Proteins and nucleic acids encoding same							
Patent: WO 0216599-A 74 28-FEB-2002;							
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)							
FEATURES							
source							
1..1183							
/organism="Homo sapiens"							
/mol_type="unassigned DNA"							
/db_xref="taxon:9606"							
ORIGIN							
Query Match 100.0%; Score 1183; DB 6; Length 1183;							
Best Local Similarity 100.0%; Pred. No.0;							
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
QY	1	ACCTAGAGCTCCAGGACCCACCGCTGTCTCTGTGACAGAGCTCAAGGCGCTGGG	60				
DB	1	ACCTAGAGCTCCAGGACCCACCGCTGTCTCTGTGACAGAGCTCAAGGCGCTGGG	60				
QY	61	CGTTCCCTCCCTGGCTGGCTGTCTGGAGGTTCCCGATCCAGATCCCTAAGGAG	120				
DB	61	CGTTCCCTCCCTGGCTGGCTGTCTGGAGGTTCCCGATCCAGATCCCTAAGGAG	120				
QY	121	CATGGGGAGCTGATCCATCCCTGGTGTAACAACTGCTGACTGCAGACAGATGCTGAGCT	180				
DB	121	CATGGGGAGCTGATCCATCCCTGGTGTAACAACTGCTGACTGCAGACAGATGCTGAGCT	180				
QY	181	ACCCAAACCAACACCTAGCTCTCCCTGAGAGTCTCCAGGCTGAGAGGTTCTGGGTG	240				
DB	181	ACCCAAACCAACACCTAGCTCTCCCTGAGAGTCTCCAGGCTGAGAGGTTCTGGGTG	240				
QY	241	TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCAAAGCCCTACCTGTCCA	300				
DB	241	TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCAAAGCCCTACCTGTCCA	300				
QY	301	GCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGGCTTTGATGACAAACCAATTC	360				
DB	301	GCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGGCTTTGATGACAAACCAATTC	360				


```
QY 901 ACTCAGGCCCTGGTGACCAATTACTCTGAGCTGGCGGATGACATCTGCTGCCTACTCAA 960
Db 283 ACTCAGGCCCTGGTGACCAATTACTCTGAGCTGGCGGATGACATCTGCTGCCTACTCAA 224
QY 961 GGAGCCCTGTCTCTCTGAGAGGGCTGGCCCTCCCTGGCAAGGATATACCCCTACTCTGT 1020
Db 223 GGAGCCCTGTCTCTCTGAGAGGGCTGGCCCTCCCTGGCAAGGATATACCCCTACTCTGT 164
QY 1021 GACTGTGAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGA 1080
Db 163 GACTGTGAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGA 104
QY 1081 AGCTGCCACAGGAGAGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTA 1140
Db 103 AGCTGCCACAGGAGAGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTA 44
QY 1141 CATCAGCCTGAATGACGAGGGTGTCTCTTTGGATGATGCCTAG 1183
Db 43 CATCAGCCTGAATGACGAGGGTGTCTCTTTGGATGATGCCTAG 1

RESULT 3
AX452880
LOCUS AX452880 2567 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent WO0242457.
ACCESSION AX452880
VERSION AX452880.1 GI:21712520
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villar, J.J. and
Kanner, S.B.
TITLE Cloning and expression of human slap-2: a novel sh2/sh3
domain-containing human slap homologue having immune cell-specific
expression
JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;
Bristol-Myers Squibb Co. (US)
FEATURES
Source 1..2567
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 95.7%; Score 1132; DB 6; Length 2567;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCTAGAGCTCCAAGGACCCCAAGGCTGTCTCTGTGACAGAGCTCAAAGGGCCCTGGG 60
Db 18 AGCTAGAGCTCCAAGGACCCCAAGGCTGTCTCTGTGACAGAGCTCAAAGGGCCCTGGG 77
QY 61 CTTTCCCTCCCTGGCTGGCTGTCTGGAGGGTTCCTCCAGTCCAGAACTCCCTAAGGAG 120
Db 78 CTTTCCCTCCCTGGCTGGCTGTCTGGAGGGTTCCTCCAGTCCAGAACTCCCTAAGGAG 137
QY 121 CATGGGGCAGCTGATCAATCCCTGGTGATCAAACTGTGCTGACTGCAGACAGATGCTGAGCT 180
Db 138 CATGGGGCAGCTGATCAATCCCTGGTGATCAAACTGTGCTGACTGCAGACAGATGCTGAGCT 197
QY 181 ACCAAACCAACACTAGCTCTCTCCCTGAAGATCCCTCCAGGCTGAGAGATTTCTGGGTG 240
Db 198 ACCAAACCAACACTAGCTCTCTCCCTGAAGATCCCTCCAGGCTGAGAGATTTCTGGGTG 257
QY 241 TCCTAGGACCAAGAGACACTGACAGCTTCCAGAGAGGGCCCAAGCCCTAACCTGCTCCA 300
Db 258 TCCTAGGACCAAGAGACACTGACAGCTTCCAGAGAGGGCCCAAGCCCTAACCTGCTCCA 317
QY 301 GCCAGAGCATGCGTCTCAGCAGAGAGTGTCTTCCCAAGCGCTTTGTATGACAAACCAATTTC 360
```

```
Db 318 GCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTTGATGACAAACCAATTTC 377
QY 361 CTCGATGATGTGCTTCTGAGTGTCTCTGAGAAACAATGGGAAGTCTCTCCAGCAGGAAG 420
Db 378 CTCGATGATGTGCTTCTGAGTGTCTCTGAGAAACAATGGGAAGTCTCTCCAGCAGGAAG 437
QY 421 AAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTCCAAAGCCAGGGACCTGTGACCAT 480
Db 438 AAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTCCAAAGCCAGGGACCTGTGACCAT 497
QY 481 GGAAGCAGAGAGAAGCAAGGCCACAGCCCTGTGGCAGTTTCCCGCAGGTGGCC 540
Db 498 GGAAGCAGAGAGAAGCAAGGCCACAGCCCTGTGGCAGTTTCCCGCAGGTGGCC 557
QY 541 GGCAGAGCTGTGCTGAGACTGGGAGGCCATTGACCATCTCTCTGAGAGTGGAGACTG 600
Db 558 GGCAGAGCTGTGCTGAGACTGGGAGGCCATTGACCATCTCTCTGAGAGTGGAGACTG 617
QY 601 GTGGACGGTGTCTGTGAAGTCTCAGGACAGAGATATAACATCCCCAGCGTCCACGTGGG 660
Db 618 GTGGACGGTGTCTGTGAAGTCTCAGGACAGAGATATAACATCCCCAGCGTCCACGTGGG 677
QY 661 CAAAGTCTCCATGGTGTGCTGTATGAGGGCTGTGAGCAGGAGAGAGAGAGAACTGCT 720
Db 678 CAAAGTCTCCATGGTGTGCTGTATGAGGGCTGTGAGCAGGAGAGAGAGAACTGCT 737
QY 721 GTTGTACCTCTGGAAACCTCTGGAGGGCCCTTCTCATCCGGAGAGCAGACAGAGAGAG 780
Db 738 GTTGTACCTCTGGAAACCTCTGGAGGGCCCTTCTCATCCGGAGAGCAGACAGAGAGAG 797
QY 781 CTCCTACTCTCTGTCTAGTCCGCTCAGCGCCCTGTGATCTGGGACCGGATCAGACACTA 840
Db 798 CTCCTACTCTCTGTCTAGTCCGCTCAGCGCCCTGTGATCTGGGACCGGATCAGACACTA 857
QY 841 CAGGATCCACTCTGTTGACAAATGGCTGTGATACATCTCACCGCGCTCACCTTCCCTC 900
Db 858 CAGGATCCACTCTGTTGACAAATGGCTGTGATACATCTCACCGCGCTCACCTTCCCTC 917
QY 901 ACTCCAGGCCCTGGTGGACCATTAATCTCTGAGCTGGCGGATGACATCTGTGCTACTCAA 960
Db 918 ACTCCAGGCCCTGGTGGACCATTAATCTCTGAGCTGGCGGATGACATCTGTGCTACTCAA 977
QY 961 GAGGCCCTGTCTCTGAGAGGGCTGGCCCTCCCTGCAAGGATATACCCCTACTCTGT 1020
Db 978 GAGGCCCTGTCTCTGAGAGGGCTGGCCCTCCCTGCAAGGATATACCCCTACTCTGT 1037
QY 1021 GACTGTGACAGAGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCTGTTTCTGA 1080
Db 1038 GACTGTGACAGAGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCTGTTTCTGA 1097
QY 1081 AGCTGCCACAGGGAGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTA 1140
Db 1098 AGCTGCCACAGGGAGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTA 1157
QY 1141 CATCAGCCTGAATGACGAGGGTGTCTTTGGATGATGCCTAG 1183
Db 1158 CATCAGCCTGAATGACGAGGGTGTCTTTGGATGATGCCTAG 1200

RESULT 4
BC042041 2538 bp mRNA linear PRI 07-OCT-2003
LOCUS BC042041
DEFINITION Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (cdna
clone MGC:49845 IMAGE:4429896), complete cds.
ACCESSION BC042041
VERSION BC042041.1 GI:27469842
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2538)
```

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huly, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 2538)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (23-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue procurement: DCTD/DRP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 88 Row: a Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28416422.

FEATURES
source
1..2538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:49845 IMAGE:4429896"
/tissue_type="Prostate, adenocarcinoma."
/clone_lib="NIH MGC_91"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2538
/gene="SLA2"
/note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845, SLAP2"
/db_xref="LocusID:84174"
/db_xref="MIM:606577"
363..1148
/codon_start=1
/product="Src-like adaptor 2, isoform a"
/protein_id="AAH42041.1"
/db_xref="GI:27469843"
/db_xref="LocusID:84174"
/translation="MGSLPSPRRKSLPSPSSSSVQGGPVPTMEERSKATAVALGSPS
AGGPAELSLRGEPLTIVSDGDWMTVLVSEVSGREYIPSVHVAKVSHGMDYEGLSRE
KAEELLPLPGNPGAFILRESQTRGSYSLSRLSPASMDIRIRHICLDNGWLYI

misc_feature
488..632
/note="SH3; Region: SH3 domain. SH3 (Src homology 3) domains are often indicative of a protein involved in signal transduction related to cytoskeletal organization. First described in the Src cytoplasmic tyrosine kinase. The structure is a partly opened beta barrel"
/db_xref="CDD:pfam00018"
642..890
/note="SH2; Region: SH2 domain"
/db_xref="CDD:pfam00017"

misc_feature
88.4%; Score 1046; DB 9; Length 2538;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN
Query Match
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 36 GTGACAGAGCTCAAGGGCCCTGGGCTTCCCTCGCTGGCTGGCTGGAGGGT 95
DB 1 GTGACAGAGCTCAAGGGCCCTGGGCTTCCCTCGCTGGCTGGCTGGAGGGT 60
QY 96 TCCCGAGTCCAGATCCCTAAGGCGATCGGCGAGCTGATCCATCCCTGCTGACAACT 155
DB 61 TCCCGAGTCCAGATCCCTAAGGCGATCGGCGAGCTGATCCATCCCTGCTGACAACT 120
QY 156 GCTGACTGCAGACAGATGCTGAGTACCCAAACCAACACCTAGCTCTCCCTGAAGATCC 215
DB 121 GCTGACTGCAGACAGATGCTGAGTACCCAAACCAACACCTAGCTCTCCCTGAAGATCC 180
QY 216 TCCCGAGTCCAGATCCCTAAGGCGATCGGCGAGCTGATCCATCCCTGCTGACAACT 275
DB 181 TCCCGAGTCCAGATCCCTAAGGCGATCGGCGAGCTGATCCATCCCTGCTGACAACT 240
QY 276 GGCCCCCAAGCCCTAACCTGTCAGGCGAGATCGCTCTCAGCAGAGCTGCTTCTCCCA 335
DB 241 GGCCCCCAAGCCCTAACCTGTCAGGCGAGATCGCTCTCAGCAGAGCTGCTTCTCCCA 300
QY 336 AGCTTTGATGACAAACCAATTCCTCGATGATGCTGCTGAGTGTCTCTGAGGAA 395
DB 301 AGCTTTGATGACAAACCAATTCCTCGATGATGCTGCTGAGTGTCTCTGAGGAA 360
QY 396 CAATGGAGTCTGCCAGCAGAGAAATCTTCCAAAGCCCAAGCTTGAAGTCTCTCTG 455
DB 361 CAATGGAGTCTGCCAGCAGAGAAATCTTCCAAAGCCCAAGCTTGAAGTCTCTCTG 420
QY 456 TCCAGGCCAGGAGCTGTGACCTGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCC 515
DB 421 TCCAGGCCAGGAGCTGTGACCTGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCC 480
QY 516 TGGGAGATTCGGGCGAGTGGCGCGGAGCTGCTGAGAGCTCCGGGAGCATTTGA 575
DB 481 TGGGAGATTCGGGCGAGTGGCGCGGAGCTGCTGAGAGCTCCGGGAGCATTTGA 540
QY 576 CCATCTCTCTGAGGATGAGAGCTGTGTGACCGGTGTGTCTGAAAGTCTCAGGAGAGT 635
DB 541 CCATCTCTCTGAGGATGAGAGCTGTGTGACCGGTGTGTCTGAAAGTCTCAGGAGAGT 600
QY 636 ATAACTCCCGAGCTCCAGCTGGGCAAGTCTCCCATGGGTGGCTGATGAGGCGCTGA 695
DB 601 ATAACTCCCGAGCTCCAGCTGGGCAAGTCTCCCATGGGTGGCTGATGAGGCGCTGA 660
QY 696 GCAGGAGAGAAAGCAGAGGAATCTGTCTTACCTGGGAACCTCGAGGGGCTTCTCTCA 755
DB 661 GCAGGAGAGAAAGCAGAGGAATCTGTCTTACCTGGGAACCTCGAGGGGCTTCTCTCA 720
QY 756 TCCGGAGAGCCAGACACAGAGAGGCTCTTACTCTGTGTCAGTCCGCTCAGCCGCTG 815
DB 721 TCCGGAGAGCCAGACACAGAGAGGCTCTTACTCTGTGTCAGTCCGCTCAGCCGCTG 780
QY 816 CATCTCGGAGCCGATCAGACACTCAGGATCCATGCTTCCAAATGGCTGGCTGTACA 875
DB 781 CATCTCGGAGCCGATCAGACACTCAGGATCCATGCTTCCAAATGGCTGGCTGTACA 840

```
QY 876 TCTCACCGGCGCTCAGCTTCCCTCACTCAGGCGCTGGTGACCAATTAATCTAGCTGG 935
Db 841 TCTCACCGGCGCTCAGCTTCCCTCACTCAGGCGCTGGTGACCAATTAATCTAGCTGG 900
QY 936 CGGATGACATCTGCTGCTCACTCAAGGAGCCCTGTGCTGACAGGGCTGGCCGCTCC 995
Db 901 CGGATGACATCTGCTGCTCACTCAAGGAGCCCTGTGCTGACAGGGCTGGCCGCTCC 960
QY 996 CTGGCAAGGATATACCCCTACCTGTGCTGACAGGAGCAACCTCAACTGGAAGAGC 1055
Db 961 CTGGCAAGGATATACCCCTACCTGTGCTGACAGGAGCAACCTCAACTGGAAGAGC 1020
QY 1056 TGGACAGCTCCCTCCCTGTTTCTGAGCTGCGCACAGGGAGGAGTCTCTCTCAGTGAGG 1115
Db 1021 TGGACAGCTCCCTCCCTGTTTCTGAGCTGCGCACAGGGAGGAGTCTCTCTCAGTGAGG 1080
QY 1116 GTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATCAGGAGGCTGTCTCTTTGGATG 1175
Db 1081 GTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATCAGGAGGCTGTCTCTTTGGATG 1140
QY 1176 ATGCTAG 1183
Db 1141 ATGCTAG 1148

RESULT 5
AX511153
LOCUS
DEFINITION
Sequence 4 from Patent WO0242452.
ACCESSION
AX511153
VERSION
AX511153.1 GI:23392046
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
McGlade, J.C. and Loreto, M.P.
Adapter gene
Patent: WO 0242452-A 4 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
Location/Qualifiers
source
1. .786
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 62.1%; Score 735; DB 6; Length 786;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGGAGTCTGCCAGCAGAGAAATCTCTCCAGGCCCAAGCTTGAGTCTCTGTC 457
Db 1 ATGGGAGTCTGCCAGCAGAGAAATCTCTCCAGGCCCAAGCTTGAGTCTCTGTC 60
QY 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG 517
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG 120
QY 518 GGCAGTTTCCCGGAGTGCCCGCCGAGTGTGCTGAGCTCGGGAGCCATTGACC 577
Db 121 GGCAGTTTCCCGGAGTGCCCGCCGAGTGTGCTGAGCTCGGGAGCCATTGACC 180
QY 578 ATCGTCTCTGAGGATGAGACTGTGTGACGCTGTCTGAAGTCTCAGGACAGAGATAT 637
Db 181 ATCGTCTCTGAGGATGAGACTGTGTGACGCTGTCTGAAGTCTCAGGACAGAGATAT 240
QY 638 AACATCCCGACGCTCCACGCTGGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 697
Db 241 AACATCCCGACGCTCCACGCTGGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300
```

```
QY 698 AGGAGAAAGCAGAGGAGTCTGTTTACCTGGAAACCTGGAGGGGCTTCTCTCATC 757
Db 301 AGGAGAAAGCAGAGGAGTCTGTTTACCTGGAAACCTGGAGGGGCTTCTCTCATC 360
QY 758 CGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTGCTCCGCTCAGCGCCGCTGCA 817
Db 361 CGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTGCTCCGCTCAGCGCCGCTGCA 420
QY 818 TCCTGGGACCGGATCAGACACTAGGATCCACTGCTGCAATGGCTGGCTGTACATC 877
Db 421 TCCTGGGACCGGATCAGACACTAGGATCCACTGCTGCAATGGCTGGCTGTACATC 480
QY 878 TCACCGCGCTCACCTTCCCTCACTCCAGGCGCTGGTGACCAATTAATCTAGCTGGCG 937
Db 481 TCACCGCGCTCACCTTCCCTCACTCCAGGCGCTGGTGACCAATTAATCTAGCTGGCG 540
QY 938 GATGACATCTGCTGCTCACTCAAGGAGCCCTGTGCTGACAGGGCTGGCCGCTCCCT 997
Db 541 GATGACATCTGCTGCTCACTCAAGGAGCCCTGTGCTGACAGGGCTGGCCGCTCCCT 600
QY 998 GSCAAGGATATACCCCTACCTGTGCTGACAGGAGCAACCTCAACTGGAAGAGCTG 1057
Db 601 GSCAAGGATATACCCCTACCTGTGCTGACAGGAGCAACCTCAACTGGAAGAGCTG 660
QY 1058 GACAGCTCCCTCTCTGTTTCTGAAGTCCACAGGGAGGAGTCTCTCTCAGTGAGGT 1117
Db 661 GACAGCTCCCTCTCTGTTTCTGAAGTCCACAGGGAGGAGTCTCTCTCAGTGAGGT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACAGGGCTGTCTCTTGGATGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACAGGGCTGTCTCTTGGATGAT 780
QY 1178 GCTAG 1183
Db 781 GCTAG 786

RESULT 6
AX572845
LOCUS
DEFINITION
Sequence 1 from Patent WO02055707.
ACCESSION
AX572845
VERSION
AX572845.1 GI:26004935
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Holland, S.J., Mendenhall, M.K., Pardo, J., Spencer, C., Fu, A.C.,
Luo, Y., Payan, D.G., Mancebo, H.S., Wu, J., Zhou, X., Shen, M.,
Liao, X.C. and Sheng, N.
Cloning of an inhibitor of antigen-receptor signaling by a
retroviral-based functional screen
Patent: WO 02055707-A 1 18-JUL-2002;
Rigel Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source
1. .786
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
1..786
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD58542.1"
/db_xref="GI:26004936"
/db_xref="REMBL:CAD58542"
/translation="MGSLPSRRKSLSPSLSSSSVQGGPVTMEERSKATAVALGSFF
AGGDELSELGEPLTVSEGDWMTVLSEVSGREYNTPSHVAVXSHGWLVEGLSRE
KAEELLLPFGGAFILRESQTRGSYSLVRLSRPASMDRIEYRHLCDNGWLYI
SPRLTFPSLQALVDHYSELADDTICLLKEPCVLQAGPLPKSDIPLVTVTPLNWK
ELDSLLFLSEATGESLSEGLRESLSFYISLNDSEAVSLDDA"
```

ORIGIN


```

Db      541  GATGACATCTGCTGCCTACTCAAGAGCCCTGTGCTCTGAGAGGGCTGGCCCGCTCCCT 600
Qy      998  GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCAACTGGAAGAGCTG 1057
Db      601  GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCAACTGGAAGAGCTG 660
Qy      1058  GACAGCTCCCTCCCTGTTTCTGAAAGCTGCCACAGGGGAGAGTCTCTTCTCAGTGAGGGT 1117
Db      661  GACAGCTCCCTCCCTGTTTCTGAAAGCTGCCACAGGGGAGAGTCTCTTCTCAGTGAGGGT 720
Qy      1118  CTCGGAGAGTCCCTCAGTTCTACATCAGCTGAATGACAGAGGCTGTCTCTTTGGATGAT 1177
Db      721  CTCGGAGAGTCCCTCAGTTCTACATCAGCTGAATGACAGAGGCTGTCTCTTTGGATGAT 780
Qy      1178  GCCTAG 1183
Db      781  GCCTAG 786

RESULT 8
AF326353
LOCUS      Homo sapiens Src-like adapter protein-2 mRNA, complete cds.
DEFINITION
ACCESSION AF326353
VERSION    AF326353.1 GI:16797891
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 786)
AUTHORS   Holland,S.J., Liao,X.C., Mendenhall,M.K., Zhou,X., Pardo,J.,
            Chu,P., Spencer,C., Fu,A.C., Sheng,N., Yu,P., Pali,E., Nagin,A.,
            Shen,M., Yu,S., Chan,E., Wu,X., Li,C., Woisetschlager,M.,
            Aversa,G., Kolbinger,F., Bennett,M.K., Molineaux,S., Luo,Y.,
            Payan,D.G., Mancebo,H.S.Y. and Wu,J.
            Functional Cloning of Src-like Adapter Protein-2 (SLAP-2), a Novel
            Inhibitor of Antigen Receptor Signaling
            J. Exp. Med. 194 (9), 1263-1276 (2001)
            21553259
            11696592
REFERENCE  2 (bases 1 to 786)
AUTHORS   Holland,S.J., Mendenhall,M.K., Zhou,X., Spencer,C., Pardo,J.,
            Fu,A.C., Sheng,N., Shen,M., Liao,C., Luo,Y., Payan,D.G.,
            Mancebo,H.S.Y. and Wu,J.
            Direct Submission
            Submitted (05-DEC-2000) Rigol Pharmaceutical Inc., 240 East Grand
            Avenue, South San Francisco, CA 94080, USA
            Location/Qualifiers
            1..786
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /chromosome="20"
            2..786
               /note="SLAP2; SH2 and SH3 domain-containing adaptor"
               /codon_start=1
               /product="Src-like adapter protein-2"
               /protein_id="AA129204.1"
               /db_xref="GI:16797892"
               /translation="MGSLPSRRKSLPSPSSVSSVQGGPVTMEARSKATAVALGSPF
               AGPAELSLRLGEPLTVISEDQDWTVLSVSGREYNIPSVHVAKVSHGLYELGLRE
               KAEELLIPGNPGAFITRESQTRGVSLSVRLSRPDRIRHRIHICLDNGWLYI
               SPLTFPSLQALNPHYSELADIDICLLKEPCVLRAGPLPKQDIPLPVTQVRFPLNWK
               ELDSLLFSEATGEESSLSEGLRESLSFYISLNDVAVSLDDA"

CDS
1..786
/codon_start=1
/product="Src-like adapter protein-2"
/protein_id="AA129204.1"
/db_xref="GI:16797892"
/translation="MGSLPSRRKSLPSPSSVSSVQGGPVTMEARSKATAVALGSPF
AGPAELSLRLGEPLTVISEDQDWTVLSVSGREYNIPSVHVAKVSHGLYELGLRE
KAEELLIPGNPGAFITRESQTRGVSLSVRLSRPDRIRHRIHICLDNGWLYI
SPLTFPSLQALNPHYSELADIDICLLKEPCVLRAGPLPKQDIPLPVTQVRFPLNWK
ELDSLLFSEATGEESSLSEGLRESLSFYISLNDVAVSLDDA"

ORIGIN
Query Match      62.1%; Score 735; DB 9; Length 786;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      398  ATGGGAAGCTGCCCCAGCAGAGAAATAATCTCTGCCAGCCCAAGCTTGAGTCTCTCTGTC 457

```

```

Db      1  ATGGGAAGCTGCCCCAGCAGAGAAATAATCTCTGCCAGCCCAAGCTTGAGTCTCTCTGTC 60
Qy      458  CAAGGCCAGGGACCTGTGACCATGAAGCAGAGAGAAAGCAAGCCACAGCCGTGGCCCTG 517
Db      61  CAAGGCCAGGGACCTGTGACCATGAAGCAGAGAGAAAGCAAGCCACAGCCGTGGCCCTG 120
Qy      518  GGCAGTTTCCCAGAGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCAATTGACC 577
Db      121  GGCAGTTTCCCAGAGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCAATTGACC 180
Qy      578  ATCGTCTCTGAGATGGAGACTGTGAGCGGTGCTCTGAAGTCTCAGCAGAGAGATAT 637
Db      181  ATCGTCTCTGAGATGGAGACTGTGAGCGGTGCTCTGAAGTCTCAGCAGAGAGATAT 240
Qy      638  AACATCCCCAGCGTCCACGTGGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGGCCCTGAGC 697
Db      241  AACATCCCCAGCGTCCACGTGGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGGCCCTGAGC 300
Qy      698  AGGAGAAAGCAGAGAACTGCTGTGTTTACCTGGGAACCTTGGAGGGGCTTCTCTCATC 757
Db      301  AGGAGAAAGCAGAGAACTGCTGTGTTTACCTGGGAACCTTGGAGGGGCTTCTCTCATC 360
Qy      758  CGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTCTAGTCCGCTCAGCCGCTGCA 817
Db      361  CGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTCTAGTCCGCTCAGCCGCTGCA 420
Qy      818  TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTTGACAATGGCTGGCTGTATCATC 877
Db      421  TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTTGACAATGGCTGGCTGTATCATC 480
Qy      878  TCACCGGGCTCACCTTCCCTCCTCAGTCCAGGCGCTGGTGAGCCATTACTCTCAGCTGAGCG 937
Db      481  TCACCGGGCTCACCTTCCCTCCTCAGTCCAGGCGCTGGTGAGCCATTACTCTCAGCTGAGCG 540
Qy      938  GATGACATCTGCTGCCTACTCAAGGAGCCCTGTGCTCTCAGAGGGCTGCGCCGCTCCCT 997
Db      541  GATGACATCTGCTGCCTACTCAAGGAGCCCTGTGCTCTCAGAGGGCTGCGCCGCTCCCT 600
Qy      998  GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCACTGGAAGAGCTG 1057
Db      601  GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCACTGGAAGAGCTG 660
Qy      1058  GACAGTCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCTCAGCTGAGGGT 1117
Db      661  GACAGTCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCTCAGTGAAGGT 720
Qy      1118  CTCGGAGAGTCCCTCAGTTCTACATCAGCTGAATGACAGAGGCTGTCTCTTTGGATGAT 1177
Db      721  CTCGGAGAGTCCCTCAGTTCTACATCAGCTGAATGACAGAGGCTGTCTCTTTGGATGAT 780
Qy      1178  GCCTAG 1183
Db      781  GCCTAG 786

RESULT 9
AK025645
LOCUS      Homo sapiens CDNA: FLJ21992 fib, clone HEP06554.
DEFINITION
ACCESSION AK025645
VERSION    AK025645.1 GI:10438227
KEYWORDS   oligo capping; fib (full insert sequence).
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS   Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
            Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
            Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
            NEDO human cDNA sequencing project
            Unpublished

```



```

627 AACATCCCGCAGCGTCCAGTGCACCAAGTCTCCATGGTGGCTGTATGAGGGCTCGAGC 686
698 AGGGAAGAAGCAGAGAACTGCTGTGTTACTGTGGAACCTCGAGGGGCTTCTCTCATC 757
687 AGGGAAGAAGCAGAGAACTGCTGTGTTACTGTGGAACCTCGAGGGGCTTCTCTCATC 746
758 CGGAGAGCCAGACAGAGAGGCTTCTTACTTCTGTGTCAGTCCGCTCAGCGCCCTGCA 817
747 CGGAGAGCCAGACAGAGAGGCTTCTTACTTCTGTGTCAGTCCGCTCAGCGCCCTGCA 806
818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACAATGGCTGCTGTACATC 877
807 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACAATGGCTGCTGTACATC 866
878 TCACCGCCCTCACCTTCCCTCTCACTCCAGGCGCTGTGTGAGCACTTACTGTAGCTGGCG 937
867 TCACCGCGCTCACCTTCCCTCTCACTCCAGGCGCTGTGTGAGCACTTACTGTAGCTGGCG 926
938 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGTCTTGCAGAGGCTTGGCCCGCTCCCT 997
927 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGTCTTGCAGAGGCTTGGCCCGCTCCCT 986
998 GGCAGGATATACCCCTACCTGTGACTGTGAGGAGACACCACTCAACTGGAAGAGCTG 1057
987 GGCAGGATATACCCCTACCTGTGACTGTGAGGAGACACCACTCAACTGGAAGAGCTG 1046
1058 GACAGCTCCCTCTCTTTTCTGAGCTGCCACAGGGAGAGTCTTCTCAGTGAAGGT 1117
1047 GACAGCTCCCTCTCTTTTCTGAGCTGCCACAGGGAGAGTCTTCTCAGTGAAGGT 1106
1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGCTGATGACGAGGCTGTCTCTTTGGATGAT 1177
1107 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGCTGATGACGAGGCTGTCTCTTTGGATGAT 1166
1178 GCCTAG 1183
1167 GCCTAG 1172

RESULT 11
AX511155
LOCUS
DEFINITION
Sequence 6 from Patent WO0242452.
ACCESSION
AX511155
VERSION
AX511155.1 GI:23392047
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
McGlade, J.C. and Loreto, M.P.
Adapter gene
Patent: WO 0242452-A 6 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
Location/Qualifiers
1..737
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 40.8%; Score 483; DB 6; Length 737;
Best Local Similarity 99.8%; Pred. No. 1.9e-268;
Matches 533; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
398 ATGGGAAGTCTCCGAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTCAAGTCTCTCTGTC 457
1 ATGGGAAGTCTCCGAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTCAAGTCTCTCTGTC 60
458 CAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGCCGACAGCCGCTGCGCTG 517

```

```

61 CAAGGCCAGGAGCCTGTGACCATGTGAAGCAGAGAGAAAGCAAGCCACAGCCGTGGCCCTG 120
518 GGCAGATTTCCCGCAGGTGGCCCGGCGGAGCTGTCTGCTGAGACTCGGGAGGCCATTGACC 577
121 GGCAGATTTCCCGCAGGTGGCCCGGCGGAGCTGTCTGCTGAGACTCGGGAGGCCATTGACC 180
578 ATGCTCTCTGAGATGAGAGTGTGTGAGCGGTCTGTCTGAAGTCTCAGGAGAGATAT 637
181 ATGCTCTCTGAGATGAGAGTGTGTGAGCGGTCTGTCTGAAGTCTCAGGAGAGATAT 240
638 AACATCCCGCAGCTCCAGCTGGGCAAGTCTCCATGGGTGGCTGTATGAGGSCCTGAGC 697
241 AACATCCCGCAGCTCCAGCTGGGCAAGTCTCCATGGGTGGCTGTATGAGGSCCTGAGC 300
698 AGGGAAGAAGCAGAGAACTGCTGTGTTACTGTGGAACCTCGAGGGGCTTCTCTCATC 757
301 AGGGAAGAAGCAGAGAACTGCTGTGTTACTGTGGAACCTCGAGGGGCTTCTCTCATC 360
758 CGGAGAGCCAGACAGGAGGCTTCTTACTTCTGTCAGTCCGCTCAGCGCCCTGCA 817
361 CGGAGAGCCAGACAGGAGGCTTCTTACTTCTGTCAGTCCGCTCAGCGCCCTGCA 420
818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCACAATGGCTGCTGTACATC 877
421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCACAATGGCTGCTGTACATC 480
878 TCACCGGCTCACCTTCCCTCTCACTCCAGGCGCTGTGTGAGCACTTACTCTGAG 931
481 TCACCGGCTCACCTTCCCTCTCACTCCAGGCGCTGTGTGAGCACTTACTCTGAG 534

RESULT 12
AP290986
LOCUS
DEFINITION
Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,
complete cds, alternatively spliced.
ACCESSION
AP290986
VERSION
AP290986.1 GI:17351922
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Loreto, M.P. and McGlade, C.J.
Cloning and characterization of human Src-like adaptor protein 2
and a novel splice isoform, SLAP-2-v
Oncogene 22 (2), 266-273 (2003)
JOURNAL
MEDLINE
PUBMED
12527895
22415750
REFERENCE
2
Loreto, M.P. and McGlade, C.J.
Direct Submission
Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
Location/Qualifiers
1..737
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1..633
/notes="SLAP-2-v"
/codon_start=1
/product="Src-like adaptor protein-2 splice isoform"
/protein_id="AAL38198.1"
/db_xref="GI:17351923"
/translation="MGSLSRRKSLPSLSVQQCGPVTMEAKSKATAVALGSP
AGPALSRLGLPTLVSEGDWTVLSEGREYNIPIFHVAKVSHQWLYEGLRE
KAEELLPNPGGAFLIRESTGRGSLSLRSPASWDRIHRIHCLDNGWLYI
SPRLTFLPSLQALVDHYSEGWPAPWQGYTPTCDCAEDTTQLERAGQLPPVF"
CDS
ORIGIN

```


Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F.,
 Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G.,
 Campoliano, A., Castle, A., Choepe, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domino, M., Doyle, N., Ferreira, P., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRoque, K., Lamasares, R., Landers, T., Lehotzky, J.,
 Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Melchior, J., Meneur, L., Mihova, T., Miranda, C., Mleaga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 27, 2000 this sequence version replaced gi:7283243.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7115

Center clone name: 712 N 14

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 125577 bases at least Q40

Consensus quality: 135703 bases at least Q30

Consensus quality: 139593 bases at least Q20

Insert size: 182000; agarose-gel

Insert size: 142233; sum-of-contigs

Quality coverage: 2.6 in Q20 bases; agarose-gel

Quality coverage: 3.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 37 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1214: contig of 1214 bp in length
 1215 1314: gap of 100 bp
 1315 2673: contig of 1359 bp in length
 2674 2773: gap of 100 bp
 2774 4520: contig of 1747 bp in length
 4521 4621: gap of 100 bp
 4621 5961: contig of 1341 bp in length
 5962 6061: gap of 100 bp
 6062 7719: contig of 1858 bp in length
 7720 7819: gap of 100 bp
 7820 9799: contig of 1980 bp in length
 9800 9899: gap of 100 bp
 9900 11434: contig of 1535 bp in length
 11435 14382: contig of 2848 bp in length
 14383 17148: gap of 100 bp
 17148 17248: contig of 2666 bp in length
 17249 17249: gap of 100 bp

17249 19239: contig of 1991 bp in length
 19240 19339: gap of 100 bp
 19340 21102: contig of 1763 bp in length
 21103 21202: gap of 100 bp
 21203 23371: contig of 2169 bp in length
 23372 23471: gap of 100 bp
 23472 25782: contig of 2311 bp in length
 25783 25882: gap of 100 bp
 25883 28924: contig of 2942 bp in length
 28925 31619: contig of 2695 bp in length
 31620 34679: gap of 100 bp
 34680 34779: contig of 2960 bp in length
 34780 37527: gap of 100 bp
 37528 37627: contig of 2748 bp in length
 37628 40146: gap of 100 bp
 40147 40246: contig of 2519 bp in length
 40247 43743: contig of 3497 bp in length
 43744 47133: gap of 100 bp
 47134 47233: contig of 3290 bp in length
 47234 51024: contig of 3790 bp in length
 51025 51124: gap of 100 bp
 51125 54935: contig of 3812 bp in length
 54936 55035: gap of 100 bp
 55036 59553: contig of 4518 bp in length
 59554 59653: gap of 100 bp
 59654 64428: contig of 4775 bp in length
 64429 64529: gap of 100 bp
 64530 69211: contig of 4683 bp in length
 69212 72901: gap of 100 bp
 72902 73001: contig of 3590 bp in length
 73002 76724: contig of 3723 bp in length
 76725 76824: gap of 100 bp
 76825 81179: contig of 4355 bp in length
 81180 81279: gap of 100 bp
 81280 87009: contig of 5730 bp in length
 87010 87109: gap of 100 bp
 87110 90855: contig of 3746 bp in length
 90856 90955: gap of 100 bp
 90956 96520: contig of 5585 bp in length
 96521 102321: gap of 100 bp
 102322 102421: contig of 5701 bp in length
 102422 108233: gap of 100 bp
 108234 108333: contig of 5872 bp in length
 108334 116690: contig of 8296 bp in length
 116691 125264: gap of 100 bp
 125265 125364: contig of 8475 bp in length
 125365 136355: gap of 100 bp
 136356 136454: contig of 10990 bp in length
 136455 145833: gap of 100 bp
 145834 145833: contig of 9379 bp in length.

FEATURES
Location/Qualifiers

1. 145833
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="20"
 /clone="RP11-712N14"
 /clone_lib="RPC1-11 Human Male BAC"
 1. 1214
 /note="assembly_fragment"
 1315. 2673
 /note="assembly_fragment"
 2774. 4520
 /note="assembly_fragment"
 4621. 5961
 /note="assembly_fragment"
 6062. 7719
 /note="assembly_fragment"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

end of the gene for a novel protein tyrosine kinase, a heterogeneous nuclear ribonucleoprotein A3 pseudogene, the gene for three isoforms of a novel protein similar to putative RAB5-interacting protein, the TGF2 gene for TGF(beta)-induced transcription factor 2 with two isoforms, the MYRL2 gene for myosin regulatory light chain 2 (smooth muscle isoform), the 3' end of the gene KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4) with two isoforms and a novel gene, complete sequence.

AL050318 GI:9581785
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL050318.13 GI:9581785
HTG; CpG island; heterogeneous ribonucleoprotein; KIAA0964; myosin regulatory light chain; MYRL2; RAB5-interacting protein; SH2 domain; TGF2; transcription factor; tyrosine kinase.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145068)
Lloyd, D.
Direct Submission
Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 28, 2000 this sequence version replaced gi:5924017.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
This sequence is the entire insert of clone RP5-977B1. The true left end of clone RP3-460J9 is at 63245 in this sequence. The true right end of clone CTD-182L9 is at 62218 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-977B1 is from the library RCI-5 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES
source
repeat_region
misc_feature
misc_feature
repeat_region
repeat_region
repeat_region

1. .145068
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP5-977B1"
/clone_lib="RCI-5"
568. .875
/note="AluX repeat: matches 1. .311 of consensus"
1024. .1580
/note="match: GSS: Em:AQ620498"
1047. .1629
/note="match: GSS: Em:AQ779162"
1631. .1923
/note="AluSg repeat: matches 1. .294 of consensus"
2776. .2842
/note="I2 repeat: matches 2357. .2423 of consensus"
3891. .3949
/note="MIR repeat: matches 34. .90 of consensus"

end of the gene for a novel protein tyrosine kinase, a heterogeneous nuclear ribonucleoprotein A3 pseudogene, the gene for three isoforms of a novel protein similar to putative RAB5-interacting protein, the TGF2 gene for TGF(beta)-induced transcription factor 2 with two isoforms, the MYRL2 gene for myosin regulatory light chain 2 (smooth muscle isoform), the 3' end of the gene KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4) with two isoforms and a novel gene, complete sequence.

AL050318 GI:9581785
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL050318.13 GI:9581785
HTG; CpG island; heterogeneous ribonucleoprotein; KIAA0964; myosin regulatory light chain; MYRL2; RAB5-interacting protein; SH2 domain; TGF2; transcription factor; tyrosine kinase.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145068)
Lloyd, D.
Direct Submission
Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 28, 2000 this sequence version replaced gi:5924017.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
This sequence is the entire insert of clone RP5-977B1. The true left end of clone RP3-460J9 is at 63245 in this sequence. The true right end of clone CTD-182L9 is at 62218 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-977B1 is from the library RCI-5 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES
source
repeat_region
misc_feature
misc_feature
repeat_region
repeat_region
repeat_region

1. .145068
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP5-977B1"
/clone_lib="RCI-5"
568. .875
/note="AluX repeat: matches 1. .311 of consensus"
1024. .1580
/note="match: GSS: Em:AQ620498"
1047. .1629
/note="match: GSS: Em:AQ779162"
1631. .1923
/note="AluSg repeat: matches 1. .294 of consensus"
2776. .2842
/note="I2 repeat: matches 2357. .2423 of consensus"
3891. .3949
/note="MIR repeat: matches 34. .90 of consensus"

```

repeat_region 4458. 4611
/notes="MIR repeat: matches 67. .231 of consensus"
repeat_region 5133. 5549
/notes="LIMD2 repeat: matches 5924. .6340 of consensus"
repeat_region 5550. 5680
/notes="FLAM_C repeat: matches 1. .132 of consensus"
repeat_region 5703. 5690
/notes="AluSx repeat: matches 120. .308 of consensus"
repeat_region 5823. 6012
/notes="LIMC/D repeat: matches 5432. .5528 of consensus"
repeat_region 6588. 6889
/notes="AluSp repeat: matches 1. .303 of consensus"
gene join(7423. 7784,9960. 10039,10306. 10394,10917. 11329,
34968. 35059,36569. 36724,37531. 39351)
/genes="KIAA0964"
mRNA join(<7423. 7784,9960. 10039,10917. 11329,34968. 35059,
36569. 36724,37531. 39351)
/genes="KIAA0964"
/product="dj977B1.6.1 (KIAA0964 (ortholog of rat
PSD-95/SAP90-associated protein 4))"
/notes="match: cDNAs: Em:AB023181"
/evidence=not experimental
join(<7423. 7784,10306. 10394,10917. 11329,34968. 35059,
36569. 36724,37531. 39351)
/genes="KIAA0964"
/product="dj977B1.6.2 (KIAA0964 (ortholog of rat
PSD-95/SAP90-associated protein 4))"
/notes="match: cDNAs: Em:AB023181"
/evidence=not experimental
join(<7423. 7784,10306. 10394,10917. 11329,34968. 35059,
36569. 36724,37531. 39351)
/genes="KIAA0964"
/product="dj977B1.6.1 (KIAA0964 (ortholog of rat
PSD-95/SAP90-associated protein 4))"
/notes="match: cDNAs: Em:U67140 Em:AB000276 Em:AB000275
Em:AB023181 Em:AF009204 Em:AB003594 Em:U67988
Em:U67137 Em:U67138 Em:U67987 Em:U67139 Em:AF131778
match: ESTs: Em:A1811867 Em:A1620816 Em:AA496750
Em:AA647532 Em:A1747584 Em:AA931146 Em:A1887761
Em:AA154844 Em:AA278057 Em:A1202736 Em:A1497978
Em:AA191427 Em:AA611645 Em:AA760107 Em:AA959160
Em:AA446940 Em:AA190916 Em:AA70912 Em:AA558202 Em:A1748854
Em:A1316293 Em:A1840413 Em:A1347137 Em:AA571014
Em:AA870776 Em:AA144662 Em:A1986170 Em:AA496805
Em:AA443364 Em:AA874407 Em:A126125 Em:A1840039
Em:AA943725 Em:AA895674 Em:A1493631 Em:AA760121
Em:AA050197 Em:AA84650 Em:T35758 Em:F06905 Em:AA433898
Em:R54251 Em:A1918247 Em:H05433 Em:W42663 Em:T75108
Em:RA824392 Em:R13930 Em:A1948469 Em:AA861813 Em:AA602005
Em:A1204639 Em:N36776 Em:F25916 Em:F12931 Em:A1936733
Em:AA126063 Em:T57999 Em:A1916248 Em:T75144 Em:W07614
Em:A1972603 Em:AA434575 Em:A1179528 Em:AA031035 Em:T08824
Em:A1343454 Em:A1128723 Em:A1500286 Em:AI201649
Em:N52262 Em:R51856 Em:R96064 Em:242146 Em:T49317
Em:N22626 Em:ID28675 Em:A1652021 Em:AA295898 Em:AL049119
Em:AA22576 Em:AA046329 Em:N45937 Em:AA577212 Em:AA610855
Em:A1017860 Em:W58510 Em:T08842 Em:A1057148 Em:N22846"
/evidence=not experimental
join(<7423. 7784,9960. 10039,10917. 11329,34968. 35059,
36569. 36724,37531. 37749)
/genes="KIAA0964"
continues in dj996C2 (AL445705) and bK2182L9 (AL390374)
match: proteins: Tr:Q9Y2H0"
/codon_start=3
/evidence=not experimental
/product="dj977B1.6.2 (KIAA0964 (ortholog of rat
PSD-95/SAP90-associated protein 4))"
/protein_id="CAC16960.1"
/db_xref="GI:11323333"
/db_xref="GOA:Q9Y2H0"
/db_xref="SWISS-PROT:Q9Y2H0"
/translation="SSCLVAYKTPPPVPPRTTTSKPFISVTVOSTESADQTYLDSQD
HSEVTPAQKRLKSLIGIQVDCIQVPVKESPEATKFSQIGVQVEDWSSVPSHMS
RDTSDTDANDSSCKSSRLPDCPTPHNSISIDAGRPQAPKIAIKRNLUSYGDNS
DPALEASSLPFPPWLETSSSPAPQACRGRDYWFLKLLQAEERLEWCQCMD
KETKNLSEVLGKVLAVGSAQLMSQFQGRGLCQNLNPDANPRPTAQDLAGE
WDLQSLSDIEMKFDLHLKANSQWLVETPEKKEKPPPPVPPKPAKSPAVSR
DKASDASDKORQEARLLAAKRAASVQNSATESADSIEIYVPEAQLRL"
repeat_region 10065. 10112
/notes="6 copies 8 mer cgcctgc 79% conserved"
repeat_region 11537. 11753
/notes="MIR repeat: matches 11. .236 of consensus"
repeat_region 11767. 11782
/notes="L2 repeat: matches 2692. .2708 of consensus"
repeat_region 11783. 12083
/notes="AluSx repeat: matches 1. .300 of consensus"
repeat_region 12084. 12121
/notes="L2 repeat: matches 2661. .2692 of consensus"
repeat_region 12631. 12714
/notes="HAL1 repeat: matches 1589. .1665 of consensus"
repeat_region 13077. 13367
/notes="AluJo repeat: matches 3. .290 of consensus"
repeat_region 13440. 13742
/notes="AluSx repeat: matches 1. .304 of consensus"
repeat_region 13932. 14197
/notes="LIMB7 repeat: matches 5915. .6173 of consensus"
repeat_region 14198. 14512
/notes="AluY repeat: matches 1. .307 of consensus"
repeat_region 14513. 14756
/notes="LIMB7 repeat: matches 5650. .5915 of consensus"
repeat_region 14789. 15705
/notes="LIMB4 repeat: matches 5162. .6185 of consensus"
repeat_region 15706. 16004
Query Match 12.8%; Score 152; DB 9; Length 145068;
Best Local Similarity 100.0%; Pred. No. 3e-76;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 778 AGGCTCTTACTCTGTGTAGTCCGCTCAGCGCCCTGCATCTGGACCGGATCAGACA 837
Db 126094 AGGCTCTTACTCTGTGTAGTCCGCTCAGCGCCCTGCATCTGGACCGGATCAGACA 126035

QY 838 CTACAGATCCACTGCTTGAACATGGCTGGTGTACATCTCACCGCGCTCACCTTCCC 897
Db 126034 CTACAGATCCACTGCTTGAACATGGCTGGTGTACATCTCACCGCGCTCACCTTCCC 125975

QY 898 CTCACATCCAGGCCCTGGTGGACCATTTCTCTG 929
Db 125974 CTCACATCCAGGCCCTGGTGGACCATTTCTCTG 125943

RESULT 16
AC026539/c 145833 bp DNA linear HTG 27-APR-2000
LOCUS Homo sapiens chromosome 20 clone RP11-712N1.4 map 20, WORKING DRAFT
DEFINITION Homo sapiens chromosome 20 clone RP11-712N1.4 map 20, WORKING DRAFT

```

```

SEQUENCE, 37 unordered pieces.
AC026539
AC026539.2 GI:7656813
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145833)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 20, clone RP11-712N14
Unpublished
2 (bases 1 to 145833)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Casale,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Labrecque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
McDrim,J., Meneus,D., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Theodore,J., Tirrell,A., Travers,W., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced GI:7283243.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7115
Center clone name: 712.N.14
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 125577 bases at least Q40
Consensus quality: 135703 bases at least Q30
Consensus quality: 139593 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 142233; sum-of-contigs
Quality coverage: 2.6 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1214: contig of 1214 bp in length
* 1215 1314: gap of 100 bp
* 1315 2673: contig of 1359 bp in length
* 2674 2773: gap of 100 bp

```

FEATURES
Source

Location/Qualifiers
1..145833
/organism="Homo sapiens"
/mol_type="genomic DNA"

2774 4520: contig of 1747 bp in length
4521 4620: gap of 100 bp
4621 5961: contig of 1341 bp in length
5962 6061: gap of 100 bp
6062 7113: contig of 1658 bp in length
7114 7819: gap of 100 bp
7820 9799: contig of 1980 bp in length
9800 9899: gap of 100 bp
9900 11434: contig of 1535 bp in length
11435 11534: gap of 100 bp
11535 14382: contig of 2848 bp in length
14383 14482: gap of 100 bp
14483 17448: contig of 2666 bp in length
17449 17448: gap of 100 bp
17450 19239: contig of 1991 bp in length
19240 19339: gap of 100 bp
19340 21102: contig of 1763 bp in length
21103 21202: gap of 100 bp
21203 23371: contig of 2189 bp in length
23372 23471: gap of 100 bp
23472 25782: contig of 2311 bp in length
25783 25824: gap of 100 bp
25825 28924: contig of 2942 bp in length
28925 31619: contig of 2635 bp in length
31620 31719: gap of 100 bp
31720 34679: contig of 2960 bp in length
34680 34779: gap of 100 bp
34780 37527: contig of 2748 bp in length
37528 37627: gap of 100 bp
37629 40146: contig of 2519 bp in length
40147 40246: gap of 100 bp
40247 43743: contig of 3497 bp in length
43744 43844: gap of 100 bp
43845 47133: contig of 3230 bp in length
47134 47233: gap of 100 bp
47234 51023: contig of 3790 bp in length
51024 51123: gap of 100 bp
51124 54935: contig of 3812 bp in length
54936 55035: gap of 100 bp
55036 59553: contig of 4518 bp in length
59554 59654: gap of 100 bp
59655 64428: contig of 4775 bp in length
64429 64528: gap of 100 bp
64529 69211: contig of 4683 bp in length
69212 69311: gap of 100 bp
69312 72901: contig of 3590 bp in length
72902 73001: gap of 100 bp
73002 76724: contig of 3723 bp in length
76725 76824: gap of 100 bp
76825 81179: contig of 4355 bp in length
81180 81279: gap of 100 bp
81280 87009: contig of 5730 bp in length
87010 87109: gap of 100 bp
87110 90855: contig of 3746 bp in length
90856 90955: gap of 100 bp
90956 96520: contig of 5565 bp in length
96521 96620: gap of 100 bp
96621 102321: contig of 5701 bp in length
102322 102421: gap of 100 bp
102422 108233: contig of 5872 bp in length
108234 108394: gap of 100 bp
108395 116689: contig of 8296 bp in length
116690 116789: gap of 100 bp
116790 125244: contig of 8475 bp in length
125245 125364: gap of 100 bp
125365 136354: contig of 10990 bp in length
136355 136454: gap of 100 bp
136455 145833: contig of 9379 bp in length.


```

Query Match      3.6%; Score 43; DB 10; Length 1321;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTTCATCCGGAGAGCCAGACCCAGGAGGCT 782
DB 613 GGAGGGGCTTCTTCATCCGGAGAGCCAGACCCAGGAGGCT 655

RESULT 19
LOCUS AX511150 1348 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1 from Patent WO0242452.
ACCESSION AX511150
VERSION AX511150.1 GI:23392044
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Mcglade, J.C. and Loreto, M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 1 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source
1..1348
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match      3.6%; Score 43; DB 6; Length 1348;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTTCATCCGGAGAGCCAGACCCAGGAGGCT 782
DB 621 GGAGGGGCTTCTTCATCCGGAGAGCCAGACCCAGGAGGCT 663

RESULT 20
LOCUS AF287467 1348 bp mRNA linear ROD 03-JUN-2002
DEFINITION Mus musculus Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF287467
VERSION AF287467.1 GI:17351918
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Loreto, M.P., Berry, D.M. and McGlade, C.J.
TITLE Functional cooperation between c-Cbl and Src-like adaptor protein 2
in the negative regulation of T-cell receptor signaling
JOURNAL Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
MEDLINE 22022020
PUBMED 12024036
REFERENCE
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
source
1..1348
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/dev_stage="day 15 embryo"
282..1061
/note="SLAP-2"
/codon_start=1

CDS
/product="Src-like adaptor protein-2"
/protein_id="AAL38196.1"
/db_xref="GI:17351919"
/translation="MGSLSSRKTSPPSSSSQDPEYPMOPERHKVTAVALGSPFA
GEQRLSLPLGELPLTIISDQMTVQSEVMSYVYAKVAGHLYGLSREK
AEELLPLNGPAGFLIRSQTRRCYSLSVRLSPASWDRHRIQRLDNGWLYIS
PLRTFSLHALVHSELDAGICCPRECVLQKGLPGKTRTPPVTVTSSLNWKK
LDRSLFLFEPASGEASLSEGLRESLSYSYLSAEDPLDDA"
misc_feature 381..551
misc_feature 552..824
polyA_signal 1301..1306
ORIGIN
Query Match      3.6%; Score 43; DB 10; Length 1348;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTTCATCCGGAGAGCCAGACCCAGGAGGCT 782
DB 621 GGAGGGGCTTCTTCATCCGGAGAGCCAGACCCAGGAGGCT 663

RESULT 21
LOCUS BC052655 1384 bp mRNA linear ROD 04-NOV-2003
DEFINITION IMAGE:30040401, complete cds.
ACCESSION BC052655
VERSION BC052655.1 GI:30851667
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Dege, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, J., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, J., Smalusz, D.E.,
Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 2238257
PUBMED 12477932
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contract: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Drs. Dennis Taub, Dan Longo (NIA, USA),
Jonathan Keller (NCI, USA)
cDNA Library Preparation: Yulan Piao and Minoru Ko (National

```

Institute on Aging, NIH: <http://lgsun.grc.nia.nih.gov/cDNA/>
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: misc_mgc@nih.gov
 Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 112 Row: e Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.

FEATURES

Location/Qualifiers
 1..1384
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6Ncr"
 /db_xref="taxon:10090"
 /clone="MGC:60811 IMAGE:30040401"
 /tissue type="Hematopoietic Stem Cell,
 (lin-/C-Kit+/Sca-1-), mouse, 10 wks"
 /clone.lib="NIA Mouse Hematopoietic Stem Cell
 (lin-/C-Kit+/Sca-1-) cDNA Library (Long)"
 /lab_host="DH10B"
 /note="Vector: pSPORT1"
 1..1384
 /gene="Slp2"
 /note="synonyms: SLAP2, SLAP-2"
 /db_xref="LocusID:77799"
 /db_xref="WCI:1925049"
 330..1109
 /codon_start=1
 /product="Slp2 protein"
 /protein_id="AAH52655.1"
 /db_xref="GI:30851668"
 /db_xref="LocusID:77799"
 /translation="MGSLSRGKTSPPSSSGDPQVSPVQPERKVTVALGSPFA
 GEQAKLSRLGEPITLISEDGDMWTVQSEVSGREYHMFSTVYAKVAGHWLYEGLSREK
 AEELLLPFGGAPLIESQTRRCYSLVRLSRPASWDRIHYRIORLDNGWLYIS
 PRLTPFSLHALVHYSELADGICPLREPCVQLKGLPGKDTPTPTVPTVTSLSNKK
 LDRSLFLFEPASGEASLLSEGLRESLSYSLSAEDPLDDA"
 432..596
 /note="SH3; Region: SH3 domain, SH3 (Src homology 3)
 domains are often indicative of a protein involved in
 signal transduction related to cytoskeletal organisation.
 First described in the Src cytoplasmic tyrosine kinase.
 The structure is a partly opened beta barrel"
 /db_xref="CDD:pfam00018"
 603..899
 /note="SH2; Region: Src homology 2 domains"
 /db_xref="CDD:cd00173"

gene

CDS

misc_feature

misc_feature

ORIGIN

Query Match 3.6%; Score 43; DB 10; Length 1384;
 Best Local Similarity 100.0%; Pred. No. 1.1e-12;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTCTATCCGGAGAGCCAGCAGGAGGCT 782

DB 669 GGAGGGGCTTCTCTATCCGGAGAGCCAGCAGGAGGCT 711

RESULT 22

AL935150/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL935150
 Mouse DNA sequence from clone RP24-482A24 on chromosome 2, complete
 sequence.
 AL935150
 AL935150.10 GI:29888604
 HTG.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 163240)
 Tracey, A.
 Direct Submission
 Submitted (15-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Apr 15, 2003 this sequence version replaced gi:29536370.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun map have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP24-482A24 is
 from a Male (C57BL/6J) mouse BAC Library VECTOR: pTAREAC1.

FEATURES

source

1..163240
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP24-482A24"
 /clone_lib="RPCL-24"

ORIGIN

Query Match 3.5%; Score 41; DB 10; Length 163240;
 Best Local Similarity 100.0%; Pred. No. 1e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTCTATCCGGAGAGCCAGCAGGAGG 780

DB 104579 GGAGGGGCTTCTCTATCCGGAGAGCCAGCAGGAGG 104539

RESULT 23

LOCUS

DEFINITION

ACCESSION

VERSION

AX452884
 Sequence 5 from Patent WO0242457.
 AX452884
 AX452884.1 GI:21712523

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
ORIGIN

synthetic construct
synthetic construct
artificial sequences.

1
Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villiar, J.J. and Kanner, S.B.
Cloning and expression of human slap-2: a novel sh2/sh3 domain-containing human slap homologue having immune cell-specific expression
Patent: WO 0242457-A 5 30-MAY-2002;
Bristol-Myers Squibb Co. (US)
Location/Qualifiers
1..30
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="GENE TRAPPER PY750 PRIMER"

Query Match 2.5%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

923 TACTCTGAGCTGGCGGATGACATCTGCTGC 952
|||||
Db 1 TACTCTGAGCTGGCGGATGACATCTGCTGC 30

RESULT 24
AC125701/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-12L23, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC125701
AC125701 GI:25198800
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 299947)
Munzy, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Louisedge, H., Lorado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwackelme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfanckoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivas, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smales, D., Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, D., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yan, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 299947)
Worley, K. C.
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 299947)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 23, 2002 this sequence version replaced gi:23673264.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDYR
Center clone name: CH230-12L23
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 219909 bases at least Q40
Consensus quality: 221402 bases at least Q30
Consensus quality: 222810 bases at least Q20
Estimated insert size: 226889; sum-of-contigs estimation
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 3989: contig of 3989 bp in length

* 3990 4089: gap of unknown length
 * 4090 293945: contig of 289856 bp in length
 * 293946 294045: gap of unknown length
 * 294046 295078: contig of 1033 bp in length
 * 295079 295178: gap of unknown length
 * 295179 299947: contig of 4769 bp in length.

FEATURES

source
 1. 299947
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-12L23"
 misc_feature
 1. 1413
 /note="wgs_end_extension"
 clone_end:"7"
 2356_3989
 /note="wgs_end_extension"
 clone_end:"7"
 misc_feature
 4090_5450
 /note="wgs_end_extension"
 clone_end:"7"
 misc_feature
 complement(5182..6062)
 /note="clone_boundary"
 clone_end:"7"
 site:ECORI
 end_sequence:BH316652"
 complement(5197..6273)
 /note="clone_boundary"
 clone_end:"7"
 site:ECORI
 end_sequence:BH316652"

ORIGIN

Query Match 2.5% Score 29; DB 2; Length 299947;
 Best Local Similarity 100.0%; Pred. No. 9.1e-05;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CTTGACATGGCTGCTGCTATCATCC 882
 |||||
 Db 133022 CTTGACATGGCTGCTGCTATCATCC 132994

RESULT 25
 AC123560
 LOCUS AC123560 322940 bp DNA linear HTG 22-SEP-2002
 DEFINITION Rattus norvegicus clone CH230-61G22, *** SEQUENCING IN PROGRESS
 ***, 12 unordered pieces.
 ACCESSION AC123560
 VERSION AC123560.3 GI:23267691
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 322940)
 Murny D, Marie, Metzker M, Lee, Abramson S, Adams C, Alder J, Allen C, Allen H, Alsbrooks S, Amin A, Anguiano D, Anyalebechi V, Aoyagi A, Ayodeji M, Baca E, Baden H, Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F, Biswal K, Blair J, Blankenburg K, Blyth P, Brown M, Bryant N, Buhay C, Burrell K, Calderon E, Cardenas J, Carter K, Cavazos I, Ceasar H, Chen A, Chacko V, Chavez D, Chen R, Chen Y, Chen Z, Chu J, Cleveland C, Cockrell R, Cox C, Coyte M, Cree A, D'Souza L, Davila M, Davis C, Davy-Carroll L, De Anda C, Dederich B, Delgado O, Denison S, Deramo C, Ding Y, Dinh H, Divya K, Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K, Egan A, Escotto M, Eugene C, Evans C, Falls T, Fan G, Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P, Fraser C, Gabisi A, Ganta R, Garcia A, Garner T, Garza M, Gebregeorgis E, Gear K, Gill R, Grady M, Guerra W, Guevara W, Gunaratne P, Haaland W, Hamill C, Hamilton C, Hamilton K,

Harvey Y, Havlak P, Hawes A, Henderson N, Hernandez J, Hernandez R, Hines S, Hladun S, Hodgson A, Hogues M, Hollins B, Howells S, Hulyk S, Hume J, Idlebird D, Jackson A, Jackson L, Jacob H, Johnson B, Johnson R, Jolivet A, Karpathy S, Kelly S, Khan Z, King L, Kovar C, Kowis C, Kraft C, Lebow H, Levan J, Lewis L, Li Z, Liu J, Liu J, Liu W, Liu Y, London P, Longacre S, Lopez J, Lorensuhewa L, Loulseghe H, Lozado R, Lu X, Ma J, Maheshwari M, Mahindartne M, Mahmoud M, Malloy K, Mangum A, Mangum B, Mapua P, Martin K, Martin R, Martinez E, Mawhinney S, McLeod M, McNeill T, Meenen E, Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S, Morgan M, Morris K, Morris S, Munidasa M, Murphy M, Nair L, Nankervis C, Neal D, Newton N, Nguyen N, Norris S, Nwankwelu O, Okwuonu G, Olarnpusagoon A, Pal S, Parks K, Pasternak S, Paul H, Perez A, Perez L, Pfannkuch C, Plopper P, Poindexter A, Popovic D, Primus E, Pul L, Fuaro M, Quiroz J, Rachlin E, Reeves K, Regier M, Reigh R, Reilly B, Reilly M, Ren Y, Reuter M, Richards S, Riggs F, Rives C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, Sanders W, Savary G, Scherer S, Scott G, Shatsman S, Shen H, Shetty J, Shvartsbeyn A, Sisson I, Sitter C, Smajs D, Sneed A, Sodergren E, Song X, Sorelle R, Soosa J, Steinle M, Strong R, Sutton A, Svatek A, Tabox P, Taylor C, Taylor F, Thomas N, Thomas S, Tingey A, Trejos Z, Usmani K, Valas R, Vera V, Villalana D, Waldron L, Walker B, Wang J, Wang Q, Wang S, Warren J, Warren R, Wei X, White F, Williams G, Willson R, Wlarczyk R, Wooden H, Worley K, Wright D, Wright R, Wu J, Yakub S, Yen J, Yoon L, Yoon V, Yu F, Zhang J, Zhou J, Zhou X, Zhao S, Dunn D, von Niederhausen A, Weiss R, Smith D, Holt R, Smith H, Weinstock G, and Gibbs R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 322940)
 Worley K.C.
 Direct Submission
 Submitted (31-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 322940)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (22-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Sep 22, 2002 this sequence version replaced gi:21908372.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled with Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GXQN
 Center clone name: CH230-61G22
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 158294 bases at least Q40
 Consensus quality: 160217 bases at least Q50
 Consensus quality: 161494 bases at least Q20
 Estimated insert size: 177812; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently

```

* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 231755: contig of 231755 bp in length
* 231756 231855: gap of unknown length
* 231856 286274: contig of 54419 bp in length
* 286275 286374: gap of unknown length
* 286375 291152: contig of 12778 bp in length
* 291153 299252: gap of unknown length
* 299253 300350: contig of 1098 bp in length
* 300351 304500: gap of unknown length
* 304501 305100: contig of 1060 bp in length
* 305101 308266: contig of 1216 bp in length
* 308267 309226: gap of unknown length
* 309227 304337: contig of 1411 bp in length
* 304338 305933: contig of 2156 bp in length
* 305934 306694: gap of unknown length
* 306695 308596: contig of 1903 bp in length
* 308597 312740: contig of 4044 bp in length
* 312741 315473: contig of 2633 bp in length
* 315474 322940: contig of 7367 bp in length.
* 322941 322940: contig of 7367 bp in length.
FEATURES
  source
    1..322940
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-61G22"
      complement(231101..231755)
      /note="clone boundary"
      clone_end:Spf
      site:EcoRI
      end_sequence:BH283473"
misc_feature
  1..322940
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-61G22"
    complement(231101..231755)
    /note="clone boundary"
    clone_end:Spf
    site:EcoRI
    end_sequence:BH283473"
ORIGIN
Query Match      2.5%; Score 29; DB 2; Length 322940;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 854 CTTGACATGCTGGCTGCTACATCTCACC 882
Db 206051 CTTGACATGCTGGCTGCTACATCTCACC 206079

RESULT 26
AX443200/c
LOCUS AX443200 26 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 141 from Patent WO02165599.
ACCESSION AX443200
VERSION AX443200.1 GI:21690595
KEYWORDS
SOURCE
  synthetic construct
  synthetic construct
  artificial sequences.
REFERENCE
  1
  AUTHORS
    Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,
    Shimkets,R.A., Spytsek,K.A., Szekeres,E.S., Tomlinson,J.B.,
    Topper,J.N. and Yang,R.B.
  TITLE
    Proteins and nucleic acids encoding same
  JOURNAL
    Patent: WO 02165599-A 141 28-FEB-2002;
    Curagen Corporation (US); COR THERAPEUTICS, INC. (US)
FEATURES
  source
    1..26
      /organism="synthetic construct"
      /mol_type="unassigned DNA"

```

```

/db_xref="taxon:32630"
/note="oligonucleotide primer"
ORIGIN
Query Match      2.2%; Score 26; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 251 AAGGACACTGGCAGACTTCGAGAGG 276
Db 26 AAGGACACTGGCAGACTTCGAGAGG 1

RESULT 27
AX511164/c
LOCUS AX511164 25 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 15 from Patent WO0242452.
ACCESSION AX511164
VERSION AX511164.1 GI:23392055
KEYWORDS
SOURCE
  synthetic construct
  synthetic construct
  artificial sequences.
REFERENCE
  1
  AUTHORS
    Mcglade,J.C. and Loreto,M.P.
  TITLE
    Adapter Gene
  JOURNAL
    Patent: WO 0242452-A 15 30-MAY-2002;
    The Hospital for Sick Children (CA)
FEATURES
  source
    1..25
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="primer"
ORIGIN
Query Match      2.1%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1159 GGCTGCTCTTTGGATGATGCCTAG 1183
Db 25 GGCTGCTCTTTGGATGATGCCTAG 1

RESULT 28
AX511165
LOCUS AX511165 25 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 16 from Patent WO0242452.
ACCESSION AX511165
VERSION AX511165.1 GI:23392056
KEYWORDS
SOURCE
  synthetic construct
  synthetic construct
  artificial sequences.
REFERENCE
  1
  AUTHORS
    Mcglade,J.C. and Loreto,M.P.
  TITLE
    Adapter Gene
  JOURNAL
    Patent: WO 0242452-A 16 30-MAY-2002;
    The Hospital for Sick Children (CA)
FEATURES
  source
    1..25
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="primer"
ORIGIN
Query Match      2.1%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 773 AGGAGAGGCTCTTACTCTCTGTGTCAG 797

```

```

Db      1 AGGAGAGGCTTACTCTGTGCG 25
|||||
AX511166      25 bp      DNA      linear      PAT 27-SEP-2002
Sequence 17 from Patent WO0242452.
ACCESSION AX511166
VERSION AX511166.1 GI:23392057
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Mcglade, J.C. and Loreto, M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 17 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source
1. .25
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="primer"
ORIGIN
Query Match 2.1%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1159 GGCTGTCTCTTTGGGATGCGCTAG 1183
|||||
Db 25 GGCTGTCTCTTTGGGATGCGCTAG 1
|||||

RESULT 30
AX511163      23 bp      DNA      linear      PAT 27-SEP-2002
LOCUS
DEFINITION Sequence 14 from Patent WO0242452.
ACCESSION AX511163
VERSION AX511163.1 GI:23392054
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Mcglade, J.C. and Loreto, M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 14 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source
1. .23
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="primer"
ORIGIN
Query Match 1.9%; Score 23; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 398 ATGGGAAGTCTCCCGACGACGAG 420
|||||
Db 1 ATGGGAAGTCTCCCGACGACGAG 23
|||||

RESULT 31
AC014511      8028 bp      DNA      linear      HTG 16-NOV-1999
LOCUS
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
ACCESSION AC014511

```

```

VERSION AC014511.1 GI:6436824
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 8028)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CD1:10210887 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. .8028
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
ORIGIN
Query Match 1.9%; Score 23; DB 2; Length 8028;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 148 TACAACTGCTGACTGCAGACAG 170
|||||
Db 2482 TACAACTGCTGACTGCAGACAG 2504
|||||

RESULT 32
AB011527      14560 bp      mRNA      linear      ROD 22-AUG-1998
LOCUS
DEFINITION Rattus norvegicus mRNA for MEGF1, complete cds.
ACCESSION AB011527
VERSION AB011527.1 GI:3449285
KEYWORDS MEGF1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (sites)
AUTHORS Nakayama, M., Nakajima, D., Nagase, T., Nomura, N., Seki, N. and
Ohara, O.
TITLE Identification of high-molecular-weight proteins with multiple
EGF-like motifs by motif-trap screening
JOURNAL Genomics 51 (1), 27-34 (1998)
MEDLINE 98360089
PUBMED 9693030
REFERENCE 2 (bases 1 to 14560)
AUTHORS Nakayama, M., Nakajima, D. and Ohara, O.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-1998) Manabu Nakayama, Kazusa DNA Research
Institute, Laboratory of DNA technology; 1532-3, Yana, Kisarazu,
Chiba 292-0812, Japan (E-mail: rmanabu@kazusa.or.jp,
Tel: +81-438-52-3915, Fax: +81-438-52-3914)
COMMENT Sequence updated (05-Aug-1998).
FEATURES
source
1. .14560
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="brain"
/clone_lib="pSPORT 1"
/dev_stage="adult"
1. .14560
/genes="MEGF1"
Gene

```


Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

Location/Qualifiers
1. 177577
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="X"
/map="15B-15B"
/clone="BACR33D04 (D1243)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in PBAC3.6)"

ORIGIN
Query Match 1.9%; Score 23; DB 3; Length 177577;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 TACAACTGCTGACTGCAGACAG 170
|||||

Db 142838 TACAACTGCTGACTGCAGACAG 142816
|||||

RESULT 34
AC018489/c
LOCUS
DEFINITION
Drosophila melanogaster, chromosome X, region 15A-15B, BAC clone BACR19G19, complete sequence.
ACCESSION
AC018489.6 GI:13384325
VERSION
HTG.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 179016)
Celiker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Blandon,R.C., Rogers,Y., An,H., Baldwin,D., Bonzon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleib,J., Paragav,V., Park,S., Patel,S., Pfeiffer,B., Schaefer,F., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Swirskas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome X, region 15A-15B
Unpublished
2 (bases 1 to 179016)
Celiker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhorff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,I., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleib,J.M., Park,S., Pfeiffer,B., Richards,S., Sethi,H., Swirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Direct Submission
Submitted (13-DEC-1999) Drosophila Genome Center, Lawrence Berkeley

Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Mar 20, 2001 this sequence version replaced gi:6984332.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

Location/Qualifiers
1. 179016
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="X"
/map="15A-15B"
/clone="BACR19G19 (D1223)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in PBAC3.6)"

ORIGIN
Query Match 1.9%; Score 23; DB 3; Length 179016;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 TACAACTGCTGACTGCAGACAG 170
|||||

Db 70703 TACAACTGCTGACTGCAGACAG 70681
|||||

RESULT 35
AC128065
LOCUS
DEFINITION
Rattus norvegicus clone CH230-466B1, WORKING DRAFT SEQUENCE.
ACCESSION
AC128065.3 GI:25007797
VERSION
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 189516)
Muzny,D.,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,B., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Surrrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisai,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.I., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulsegged,H., Lozado,R.J., Lu,X., Ma,J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidaa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakolamen, O., Okwuonu, G., Olampunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Strong, R., Sutton, A., Svatek, A., Tabot, P., Taylor, C., Teilmle, M., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczyk, R., Wood, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL
 REFERENCE
 2 (bases 1 to 189516)

Direct Submission
 Worley, K.C.

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 189516)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 15, 2002 this sequence version replaced gi:23908438. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GZRY

Center clone name: CH230-466B1

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 181436 bases at least Q40

Consensus quality: 183164 bases at least Q30

Consensus quality: 184279 bases at least Q20

Estimated insert size: 188543; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 1 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

source

1. 189516
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-466B1"

misc_feature

1. 3015
 /notes="wgs end extension
 clone_end:Sp6"
 15984-43929
 /notes="clone boundary
 clone_end:Sp6"

misc_feature

end sequence:BZ214020"
 171870-171921
 /notes="clone boundary
 clone_end:T7"
 end sequence:BZ214019"

ORIGIN

Query Match 1.9% Score 23; DB 2; Length 189516;
 Best Local Similarity 100.0%; Pred.No. 0.29;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 305 GAGCATGCGTCTCAGCAGAGCTG 327
 Db 33825 GAGCATGCGTCTCAGCAGAGCTG 33847

RESULT 36

AC133702/c

LOCUS

DEFINITION

AC133702

AC133702.3

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Rattus norvegicus

Rattus norvegicus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 270729)

Muzny, D. Marie; Metzker, M. Lee; Abramson, S.; Adams, C.; Alder, J.;

Allen, C.; Allen, H.; Alsbrooks, S.; Amin, A.; Anguiano, D.;

Anyalebechi, V.; Aoyagi, A.; Ayodeji, M.; Baca, E.; Baden, H.;

Baldwin, D.; Bandaranaike, D.; Barber, M.; Barnstead, M.; Benahmed, F.;

Biswal, K.; Blair, J.; Blankenburg, K.; Blyth, P.; Brown, M.;

Bryant, N.; Buhay, C.; Burch, P.; Burrell, K.; Calderon, E.;

Cardenas, V.; Carter, K.; Cavazos, I.; Ceasar, H.; Center, A.;

Chacko, J.; Chavez, D.; Chen, R.; Chen, Y.; Chen, Y.; Chu, J.;

Cleveland, C.; Cockrell, R.; Cox, C.; Coyle, M.; Cree, A.; D'Souza, L.;

Davila, M.L.; Davis, C.; Davy-Carroll, L.; De Anda, C.; Dederich, D.;

Delgado, O.; Denson, S.; Deramo, C.; Ding, Y.; Dinh, H.; Divya, K.;

Draper, H.; Dugan-Rocha, S.; Dunn, A.; Durbin, K.; Duval, B.; Eaves, K.;

Egan, A.; Escotto, M.; Eugene, C.; Evans, C.A.; Falls, T.; Fan, G.;

Fernandez, S.; Finley, M.; Flagg, N.; Forbes, L.; Foster, M.; Foster, P.;

Fraser, C.M.; Gabisi, A.; Ganta, R.; Garcia, A.; Garner, T.; Garza, M.;

Gregg, P.; Haaland, M.; Hamill, C.; Hamilton, C.; Hamilton, C.;

Gunaratne, P.; Haavik, P.; Hawes, A.; Henderson, N.; Hernandez, J.;

Harvey, Y.; Haylak, P.; Hines, S.; Hladun, S.L.; Hodgson, A.; Hoggues, M.;

Hollins, B.; Howells, S.; Hulyk, S.; Hume, J.; Idlebird, D.; Jackson, A.;

Jackson, L.; Jacob, L.; Jiang, H.; Johnson, B.; Johnson, R.; Jolivet, A.;

Karpas, S.; Kelly, S.; Kelly, S.; Khan, Z.; King, L.; Kovar, C.;

Kowis, C.; Kraft, C.L.; Lebow, H.; Levan, J.; Lewis, L.; Li, Z.; Liu, J.;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 270729)

Muzny, D. Marie; Metzker, M. Lee; Abramson, S.; Adams, C.; Alder, J.;

Allen, C.; Allen, H.; Alsbrooks, S.; Amin, A.; Anguiano, D.;

Anyalebechi, V.; Aoyagi, A.; Ayodeji, M.; Baca, E.; Baden, H.;

Baldwin, D.; Bandaranaike, D.; Barber, M.; Barnstead, M.; Benahmed, F.;

Biswal, K.; Blair, J.; Blankenburg, K.; Blyth, P.; Brown, M.;

Bryant, N.; Buhay, C.; Burch, P.; Burrell, K.; Calderon, E.;

Cardenas, V.; Carter, K.; Cavazos, I.; Ceasar, H.; Center, A.;

Chacko, J.; Chavez, D.; Chen, R.; Chen, Y.; Chen, Y.; Chu, J.;

Cleveland, C.; Cockrell, R.; Cox, C.; Coyle, M.; Cree, A.; D'Souza, L.;

Davila, M.L.; Davis, C.; Davy-Carroll, L.; De Anda, C.; Dederich, D.;

Delgado, O.; Denson, S.; Deramo, C.; Ding, Y.; Dinh, H.; Divya, K.;

Draper, H.; Dugan-Rocha, S.; Dunn, A.; Durbin, K.; Duval, B.; Eaves, K.;

Egan, A.; Escotto, M.; Eugene, C.; Evans, C.A.; Falls, T.; Fan, G.;

Fernandez, S.; Finley, M.; Flagg, N.; Forbes, L.; Foster, M.; Foster, P.;

Fraser, C.M.; Gabisi, A.; Ganta, R.; Garcia, A.; Garner, T.; Garza, M.;

Gregg, P.; Haaland, M.; Hamill, C.; Hamilton, C.; Hamilton, C.;

Gunaratne, P.; Haavik, P.; Hawes, A.; Henderson, N.; Hernandez, J.;

Harvey, Y.; Haylak, P.; Hines, S.; Hladun, S.L.; Hodgson, A.; Hoggues, M.;

Hollins, B.; Howells, S.; Hulyk, S.; Hume, J.; Idlebird, D.; Jackson, A.;

Jackson, L.; Jacob, L.; Jiang, H.; Johnson, B.; Johnson, R.; Jolivet, A.;

Karpas, S.; Kelly, S.; Kelly, S.; Khan, Z.; King, L.; Kovar, C.;

Kowis, C.; Kraft, C.L.; Lebow, H.; Levan, J.; Lewis, L.; Li, Z.; Liu, J.;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 270729)

Muzny, D. Marie; Metzker, M. Lee; Abramson, S.; Adams, C.; Alder, J.;

Allen, C.; Allen, H.; Alsbrooks, S.; Amin, A.; Anguiano, D.;

Anyalebechi, V.; Aoyagi, A.; Ayodeji, M.; Baca, E.; Baden, H.;

Baldwin, D.; Bandaranaike, D.; Barber, M.; Barnstead, M.; Benahmed, F.;

Biswal, K.; Blair, J.; Blankenburg, K.; Blyth, P.; Brown, M.;

Bryant, N.; Buhay, C.; Burch, P.; Burrell, K.; Calderon, E.;

Cardenas, V.; Carter, K.; Cavazos, I.; Ceasar, H.; Center, A.;

Chacko, J.; Chavez, D.; Chen, R.; Chen, Y.; Chen, Y.; Chu, J.;

Cleveland, C.; Cockrell, R.; Cox, C.; Coyle, M.; Cree, A.; D'Souza, L.;

Davila, M.L.; Davis, C.; Davy-Carroll, L.; De Anda, C.; Dederich, D.;

Delgado, O.; Denson, S.; Deramo, C.; Ding, Y.; Dinh, H.; Divya, K.;

Draper, H.; Dugan-Rocha, S.; Dunn, A.; Durbin, K.; Duval, B.; Eaves, K.;

Egan, A.; Escotto, M.; Eugene, C.; Evans, C.A.; Falls, T.; Fan, G.;

Fernandez, S.; Finley, M.; Flagg, N.; Forbes, L.; Foster, M.; Foster, P.;

Fraser, C.M.; Gabisi, A.; Ganta, R.; Garcia, A.; Garner, T.; Garza, M.;

Gregg, P.; Haaland, M.; Hamill, C.; Hamilton, C.; Hamilton, C.;

Gunaratne, P.; Haavik, P.; Hawes, A.; Henderson, N.; Hernandez, J.;

Harvey, Y.; Haylak, P.; Hines, S.; Hladun, S.L.; Hodgson, A.; Hoggues, M.;

Hollins, B.; Howells, S.; Hulyk, S.; Hume, J.; Idlebird, D.; Jackson, A.;

Jackson, L.; Jacob, L.; Jiang, H.; Johnson, B.; Johnson, R.; Jolivet, A.;

Karpas, S.; Kelly, S.; Kelly, S.; Khan, Z.; King, L.; Kovar, C.;

Kowis, C.; Kraft, C.L.; Lebow, H.; Levan, J.; Lewis, L.; Li, Z.; Liu, J.;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 270729)

Muzny, D. Marie; Metzker, M. Lee; Abramson, S.; Adams, C.; Alder, J.;

Allen, C.; Allen, H.; Alsbrooks, S.; Amin, A.; Anguiano, D.;

Anyalebechi, V.; Aoyagi, A.; Ayodeji, M.; Baca, E.; Baden, H.;

Baldwin, D.; Bandaranaike, D.; Barber, M.; Barnstead, M.; Benahmed, F.;

Biswal, K.; Blair, J.; Blankenburg, K.; Blyth, P.; Brown, M.;

Bryant, N.; Buhay, C.; Burch, P.; Burrell, K.; Calderon, E.;

Cardenas, V.; Carter, K.; Cavazos, I.; Ceasar, H.; Center, A.;

Chacko, J.; Chavez, D.; Chen, R.; Chen, Y.; Chen, Y.; Chu, J.;

Cleveland, C.; Cockrell, R.; Cox, C.; Coyle, M.; Cree, A.; D'Souza, L.;

Davila, M.L.; Davis, C.; Davy-Carroll, L.; De Anda, C.; Dederich, D.;

Delgado, O.; Denson, S.; Deramo, C.; Ding, Y.; Dinh, H.; Divya, K.;

Draper, H.; Dugan-Rocha, S.; Dunn, A.; Durbin, K.; Duval, B.; Eaves, K.;

Egan, A.; Escotto, M.; Eugene, C.; Evans, C.A.; Falls, T.; Fan, G.;

Fernandez, S.; Finley, M.; Flagg, N.; Forbes, L.; Foster, M.; Foster, P.;

Fraser, C.M.; Gabisi, A.; Ganta, R.; Garcia, A.; Garner, T.; Garza, M.;

Gregg, P.; Haaland, M.; Hamill, C.; Hamilton, C.; Hamilton, C.;

Gunaratne, P.; Haavik, P.; Hawes, A.; Henderson, N.; Hernandez, J.;

Harvey, Y.; Haylak, P.; Hines, S.; Hladun, S.L.; Hodgson, A.; Hoggues, M.;

Hollins, B.; Howells, S.; Hulyk, S.; Hume, J.; Idlebird, D.; Jackson, A.;

Jackson, L.; Jacob, L.; Jiang, H.; Johnson, B.; Johnson, R.; Jolivet, A.;

Karpas, S.; Kelly, S.; Kelly, S.; Khan, Z.; King, L.; Kovar, C.;

Kowis, C.; Kraft, C.L.; Lebow, H.; Levan, J.; Lewis, L.; Li, Z.; Liu, J.;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 270729)

Muzny, D. Marie; Metzker, M. Lee; Abramson, S.; Adams, C.; Alder, J.;

Allen, C.; Allen, H.; Alsbrooks, S.; Amin, A.; Anguiano, D.;

Anyalebechi, V.; Aoyagi, A.; Ayodeji, M.; Baca, E.; Baden, H.;

Baldwin, D.; Bandaranaike, D.; Barber, M.; Barnstead, M.; Benahmed, F.;

Biswal, K.; Blair, J.; Blankenburg, K.; Blyth, P.; Brown, M.;

Bryant, N.; Buhay, C.; Burch, P.; Burrell, K.; Calderon, E.;

Cardenas, V.; Carter, K.; Cavazos, I.; Ceasar, H.; Center, A.;

Chacko, J.; Chavez, D.; Chen, R.; Chen, Y.; Chen, Y.; Chu, J.;

Cleveland, C.; Cockrell, R.; Cox, C.; Coyle, M.; Cree, A.; D'Souza, L.;

Davila, M.L.; Davis, C.; Davy-Carroll, L.; De Anda, C.; Dederich, D.;

Delgado, O.; Denson, S.; Deramo, C.; Ding, Y.; Dinh, H.; Divya, K.;

Draper, H.; Dugan-Rocha, S.; Dunn, A.; Durbin, K.; Duval, B.; Eaves, K.;

Egan, A.; Escotto, M.; Eugene, C.; Evans, C.A.; Falls, T.; Fan, G.;

Fernandez, S.; Finley, M.; Flagg, N.; Forbes, L.; Foster, M.; Foster, P.;

Fraser, C.M.; Gabisi, A.; Ganta, R.; Garcia, A.; Garner, T.; Garza, M.;

Gregg, P.; Haaland, M.; Hamill, C.; Hamilton, C.; Hamilton, C.;

Gunaratne, P.; Haavik, P.; Hawes, A.; Henderson, N.; Hernandez, J.;

Harvey, Y.; Haylak, P.; Hines, S.; Hladun, S.L.; Hodgson, A.; Hoggues, M.;

Hollins, B.; Howells, S.; Hulyk, S.; Hume, J.; Idlebird, D.; Jackson, A.;

Jackson, L.; Jacob, L.; Jiang, H.; Johnson, B.; Johnson, R.; Jolivet, A.;

Karpas, S.; Kelly, S.; Kelly, S.; Khan, Z.; King, L.; Kovar, C.;

Kowis, C.; Kraft, C.L.; Lebow, H.; Levan, J.; Lewis, L.; Li, Z.; Liu, J.;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 270729)

Muzny, D. Marie; Metzker, M. Lee; Abramson, S.; Adams, C.; Alder, J.;

Allen, C.; Allen, H.; Alsbrooks, S.; Amin, A.; Anguiano, D.;

Anyalebechi, V.; Aoyagi, A.; Ayodeji, M.; Baca, E.; Baden, H.;

Baldwin, D.; Bandaranaike, D.; Barber, M.; Barnstead, M.; Benahmed, F.;

Biswal, K.; Blair, J.; Blankenburg, K.; Blyth, P.; Brown, M.;

Bryant, N.; Buhay, C.; Burch, P.; Burrell, K.; Calderon, E.;

Cardenas, V.; Carter, K.; Cavazos, I.; Ceasar, H.; Center, A.;

Chacko, J.; Chavez, D.; Chen, R.; Chen, Y.; Chen, Y.; Chu, J.;

Cleveland, C.; Cockrell, R.; Cox, C.; Coyle, M.; Cree, A.; D'Souza, L.;

Davila, M.L.; Davis, C.; Davy-Carroll, L.; De Anda, C.; Dederich, D.;

Delgado, O.; Denson, S.; Deramo, C.; Ding, Y.; Dinh, H.; Divya, K.;

Draper, H.; Dugan-Rocha, S.; Dunn, A.; Durbin, K.; Duval, B.; Eaves, K.;

Egan, A.; Escotto, M.; Eugene, C.; Evans, C.A.; Falls, T.; Fan, G.;

Fernandez, S.; Finley, M.; Flagg, N.; Forbes, L.; Foster, M.; Foster, P.;

Fraser, C.M.; Gabisi, A.; Ganta, R.; Garcia, A.; Garner, T.; Garza, M.;

Gregg, P.; Haaland, M.; Hamill, C.; Hamilton, C.; Hamilton, C.;

Gunaratne, P.; Haavik, P.; Hawes, A.; Henderson, N.; Hernandez, J.;

Harvey, Y.; Haylak, P.; Hines, S.; Hladun, S.L.; Hodgson, A.; Hoggues, M.;

Hollins, B.; Howells, S.; Hulyk, S.; Hume, J.; Idlebird, D.; Jackson, A.;

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshe, L., Louised, H., Lozano, R., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwakoleneh, O., Okwunonu, G., Olarunpungoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, P., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, I., Rojbas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Woodin, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

Genome Center
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GEUB
 Center clone name: CH230-89p1
 Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 239257 bases at least Q40
 Consensus quality: 242117 bases at least Q30
 Consensus quality: 243763 bases at least Q20
 Estimated insert size: 244719; sum-of-coverage estimation
 Quality coverage: 7x in Q20 bases; sum-of-coverage estimation
 NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 7 contigs. The true order of the pieces

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 262132: contig of 262132 bp in length
 262133 262232: gap of unknown length
 262233 263473: contig of 1241 bp in length
 263474 263573: gap of unknown length
 263574 264640: contig of 1067 bp in length
 264641 265825: contig of 1085 bp in length
 265826 267091: contig of 1166 bp in length
 267092 267191: gap of unknown length
 267192 269114: contig of 1923 bp in length
 269115 270729: contig of 1515 bp in length.

Location/Qualifiers
 1. .270729
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-89p1"
 1. .1038
 /note="wgs_contig"
 4252. .5189
 /note="wgs_contig"
 259637. .262132
 /note="wgs_contig"

misc_feature
 misc_feature
 misc_feature

ORIGIN
 Query Match 1.9%; Score 23; DB 2; Length 270729;
 Best Local Similarity 100.0%; Pred. No. 0.28;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 GAGCATGCTCTCAGCAGAGCTG 327
 Db 45272 GAGCATGCTCTCAGCAGAGCTG 45250

RESULT 37
 AE003503/c
 LOCUS
 DEFINITION
 complete sequence.
 ACCESSION
 AE003503 AE002593 AE014298
 VERSION
 AE003503.3 GI:22832390
 KEYWORDS
 SOURCE
 ORGANISM
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 300469)
 Adams, M.D., Gelniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D.,
 Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galie, R.F.,
 George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N.,
 Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X.,
 Wan, K.H., Doyle, C., Baxter, E.G., Heit, G., Nelson, C.R., Gabor, G.L.,
 Abril, J.F., Agbayani, A., An, H.J., Andrews-Pfannkoch, C., Baldwin, D.,
 Ballew, R.M., Basu, A., Baxendale, J., Bayraktoglu, L., Beasley, E.M.,
 Beeson, K.Y., Benos, P.V., Bernier, B.P., Bhandari, D., Bolshakov, S.,
 Borkova, D., Botchan, M.R., Bouck, J., Brockstein, P., Brothier, P.,
 Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A.,
 Chandrasekhar, S., Chervak, G.N., Cawley, S., Dahlke, C., Davenport, L.B.,
 Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I.,
 Dietz, S.M., Dodson, K., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C.,
 Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C.,
 Ferreira, S., Fleischmann, W., Fosler, C., Gabrielian, A.E., Garg, N.S.,
 Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z.,

Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G.H., Ke, Z., Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A.A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusken, D.R., Pacleb, J.M., Palazolo, M., Pittman, G.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Svirska, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A., Weinstein, G.M., Weissenbach, J., Williams, S.M., Woodage, W., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, M., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.

The genome sequence of *Drosophila melanogaster*
 Science 287 (5461), 2185-2195 (2000)

20196006
 10731132

2 (bases 1 to 300469)
 Celniker, S.E., Adams, M.D., Kronmiller, B., Wan, K.H., Holt, R.A., Evans, C.A., Cocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., Barzon, J., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champe, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferreira, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Paaleb, J., Parasag, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirska, R., Tector, C., Tyler, D., Williams, S.M., Zaveri, J.S., Smith, H.O., Venter, J.C. and Rubin, G.M.

Sequencing of *Drosophila melanogaster* genome
 Unpublished

3 (bases 1 to 300469)
 Mirza, S., Crosby, M.A., Matthews, B.B., Bayraktaroglu, L., Campbell, K., Hradecky, P., Huang, Y., Kaminker, J.S., Prochnik, S.E., Smith, C.D., Tupy, J.L., Bergman, C.M., Berman, B.P., Carlson, J.W., Celniker, S.E., Clamp, M.E., Drysdale, R.A., Emmert, D., Frise, E., de Grey, A.D.N.G.J., Harris, N.L., Kronmiller, B., Marshall, B., Millburn, G.H., Richter, J., Russo, S., Searle, S.M.J., Smith, E., Shu, S., Smutniak, F., Whitfield, E.J., Ashburner, M., Gelbart, W.M., Rubin, G.M., Mungall, C.J. and Lewis, S.E.

Annotation of *Drosophila melanogaster* genome
 Unpublished

4 (bases 1 to 300469)
 Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.

Direct Submission
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

5 (bases 1 to 300469)
 FlyBase
 Direct Submission
 Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA

6 (bases 1 to 300469)
 FlyBase
 Direct Submission
 Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA

On Sep 13, 2002 this sequence version replaced gi:10728292.
 Location/Qualifiers
 1..300469
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"

/db xref="taxon:7227"
 /chromosome="X"
 /notes="genotype: y[1]; cn[1] bw[1] sp[1]; Rh6[1]"
 420..1472
 /gene="Rps19"
 /locus_tag="CG4464"
 /notes="last curated on Mon Jul 08 10:35:28 PDT 2002"
 /map="14F4-14F4"
 /db xref="FLYBASE:FBgn0010412"
 join(420..468,774..844,947..1472)
 /gene="Rps19"
 /locus_tag="CG4464"
 /product="CG4464-RA"
 /db xref="FLYBASE:FBgn0010412"
 join(439..490,774..844,947..1472)
 /gene="Rps19"
 /locus_tag="CG4464"
 /product="CG4464-RB"
 /db xref="FLYBASE:FBgn0010412"
 join(768..844,947..1472)
 /gene="Rps19"
 /locus_tag="CG4464"
 /product="CG4464-RC"
 /db xref="FLYBASE:FBgn0010412"
 join(774..844,947..1346)
 /gene="Rps19"
 /locus_tag="CG4464"
 /notes="RPS19 gene product from transcript CG4464-RA"
 /codon_start=1
 /product="CG4464-PA"
 /protein_id="AAF48633.1"
 /db xref="GI:7293252"
 /db xref="FLYBASE:FBgn0010412"
 /translation="MPGVTVKDIQDHAVTKAVAVFLKTKGLKVPDQMDIVKTAKFKE
 LAPYDPDWYVRCASILRLHYRSPAGVSGITKIYGRKRGVHSHFCRAADGAARK
 ALQALEHARLVEKHPDGGKLSIGQDLRIANQIVFKORDAAKQTGPVIVISK"
 join(774..844,947..1346)
 /gene="RPS19"
 /locus_tag="CG4464"
 /notes="RPS19 gene product from transcript CG4464-RB"
 /codon_start=1
 /product="CG4464-PB"
 /protein_id="AAN09412.1"
 /db xref="GI:22832391"
 /db xref="FLYBASE:FBgn0010412"
 /translation="MPGVTVKDIQDHAVTKAVAVFLKTKGLKVPDQMDIVKTAKFKE
 LAPYDPDWYVRCASILRLHYRSPAGVSGITKIYGRKRGVHSHFCRAADGAARK
 ALQALEHARLVEKHPDGGKLSIGQDLRIANQIVFKORDAAKQTGPVIVISK"
 join(774..844,947..1346)
 /gene="RPS19"
 /locus_tag="CG4464"
 /notes="RPS19 gene product from transcript CG4464-RC"
 /codon_start=1
 /product="CG4464-PC"
 /protein_id="AAN09413.1"
 /db xref="GI:22832392"
 /db xref="FLYBASE:FBgn0010412"
 /translation="MPGVTVKDIQDHAVTKAVAVFLKTKGLKVPDQMDIVKTAKFKE
 LAPYDPDWYVRCASILRLHYRSPAGVSGITKIYGRKRGVHSHFCRAADGAARK
 ALQALEHARLVEKHPDGGKLSIGQDLRIANQIVFKORDAAKQTGPVIVISK"
 complement(2340..2954)
 /locus_tag="CG3742"
 /notes="last curated on Fri Jul 19 18:43:02 PDT 2002"
 /map="14F4-14F4"
 /db xref="FLYBASE:FBgn0030765"
 complement(join(2340..2571,2634..2781,2847..2954))
 /locus_tag="CG3742"
 /product="CG3742-RA"
 /db xref="CG3742"
 /db xref="FLYBASE:FBgn0030765"
 complement(join(2521..2571,2634..2781,2847..2878))
 /locus_tag="CG3742"
 /note="CG3742 gene product"
 /codon_start=1

Thu Mar 25 10:05:26 2004

us-09-939-853a-74.olig.rge

```

/product="CG9742-PA"
/protein_id="AAF48634.1"
/db_xref="GI:7293253"
/db_xref="FLYBASE:FBgn0030765"
/translation="MSKAHPPEVKYMDKRWMLKNGRAVGTILRGDFPMNVLLDD
3446..12238
/locus_tag="CG4521"
/note="Last curated on Mon Jul 08 10:36:01 PDT 2002"
/map="14P4-14P5"
/db_xref="FLYBASE:FBgn0030766"
Join(3446..6050,10310..12238)
/locus_tag="CG4521"
/product="CG4521-PA"
/db_xref="FLYBASE:FBgn0030766"
4015..6048
/locus_tag="CG4521"
/note="CG4521 gene product"
/codon_start=1
/product="CG4521-PA"
/protein_id="AAF48635.1"
/db_xref="GI:7293254"
/db_xref="FLYBASE:FBgn0030766"
/translation="MDSRSSEAASSNQFIRPCGLLTIVLLQTLVMSLAIRMSPP
FAAPRPSPPTVKLNKCHSGEYINDGTGAGSEALWLPVYLQQRFEPFHGASP
Query Match 1.9%; Score 23; DB 3; Length 300469;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 TACAACTGCTGACGACGACG 170
Db 223650 TACAACTGCTGACGACGACG 223628

RESULT 38
AX443201
LOCUS 22 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 142 from Patent WO0216599.
ACCESSION AX443201
VERSION AX443201.1 GI:21690596
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,
Shinkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,
Topper,J.N. and Yang,R.B.
Proteins and nucleic acids encoding same
Patent: WO 0216599-A 142 28-FEB-2002.
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)
FEATURES
source
1..22
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="oligonucleotide primer"

ORIGIN
Query Match 1.9%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 TGAGAGATCTCGGTGCTCTA 245
Db 1 TGAGAGATCTCGGTGCTCTA 22

RESULT 39
HSPROT04/c
LOCUS 475 bp DNA linear PRI 02-FEB-1996
DEFINITION Human P protein (P) gene, exon 4.

```

```

ACCESSION U19155
VERSION U19155.1 GI:773304
KEYWORDS tyrosinase-positive oculocutaneous albinism; OCA2.
SEGMENT 4 of 25
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 475)
Lee,S.T., Nicholls,R.D., Jong,M.T., Fukai,K. and Spritz,R.A.
Organization and sequence of the human P gene and identification of
a new family of transport proteins
Genomics 26 (2), 354-363 (1995)
JOURNAL 9534928
MEDLINE 7601462
PUBMED
REFERENCE 2 (sites)
AUTHORS Rinchik,E.M., Bultman,S.J., Horsthemke,B., Lee,S.T., Strunk,K.M.,
Spritz,R.A., Avidano,K.M., Jong,M.T. and Nicholls,R.D.
A gene for the mouse pink-eyed dilution locus and for human type II
oculocutaneous albinism
Nature 361 (6407), 72-76 (1993)
JOURNAL 8421497
MEDLINE 93133287
PUBMED
REFERENCE 3 (bases 1 to 475)
AUTHORS Spritz,R.A.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1994) Richard A. Spritz, Medical Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
1..475
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q11.2-q12"
/size="179..367"
/gene="P"
/number=4
/evidence=experimental
exon
ORIGIN
Query Match 1.9%; Score 22; DB 9; Length 475;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 TCAGCAGAGCTGCTTCCCAAG 337
Db 250 TCAGCAGAGCTGCTTCCCAAG 229

RESULT 40
LOCUS 3070 bp DNA linear PAT 18-JUN-2001
DEFINITION Method for distinguishing melanocyte.
ACCESSION E24027
VERSION E24027.1 GI:13024614
KEYWORDS JP 1999103864-A/1.
SOURCE unidentified
ORGANISM unclassified.
1 (bases 1 to 3070)
Kaku,S. and Hiroharu,Y.
AUTHORS Method for distinguishing melanocyte
TITLE Patent: JP 1999103864-A 1 20-APR-1999;
JOURNAL POLA CHEM IND INC
COMMENT OS Unidentified
PN JP 1999103864-A/1
PD 20-APR-1999
PF 02-OCT-1997 JP 1997286143
PR
PI KAKU SUZUKI,HIROHARU YOKOVAMA
PC C12N15/09,A61B5/00,C12Q1/68,G01N33/15,G01N33/50,C12N15/00 CC
Strandedness: Single;

```

```

CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..3070
FT /organism='Unidentified'.
FEATURES
    source
        1..3070
            /organism='Unidentified'
            /mol_type='genomic DNA'
            /db_xref='taxon:32644'
ORIGIN
    Query Match 1..9%; Score 22; DB 6; Length 3070;
    Best Local Similarity 100.0%; Pred. No. 1.7;
    Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 316 TCAGCAGAGCTGCTTCCCAAG 337
DB 450 TCAGCAGAGCTGCTTCCCAAG 429
RESULT 41
E24028/c
LOCUS
DEFINITION Method for distinguishing melanocyte.
ACCESSION E24028
VERSION E24028.1 GI:13024615
KEYWORDS JP 1999103864-A/2.
SOURCE unclassified
ORGANISM unclassified
REFERENCE
    1 (bases 1 to 3070)
    Kaku, S. and Hirohara, Y.
    Method for distinguishing melanocyte
    Patent: JP 1999103864-A 2 20-APR-1999;
    JOURNAL POLA CHEM IND INC
COMMENT
    OS Unidentified
    PN JP 1999103864-A/2
    PD 20-APR-1999
    PF 02-OCT-1997 JP 1997286143
    PR
    PI KAKU SUZUKI, HIROHARU YOKOYAMA
    PC C12N15/09;A61B5/00;C12Q1/68;G01N33/15;G01N33/50;C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..3070
FT /organism='Unidentified'.
FEATURES
    source
        1..3070
            /organism='Unidentified'
            /mol_type='genomic RNA'
            /db_xref='taxon:32644'
ORIGIN
    Query Match 1..9%; Score 22; DB 6; Length 3070;
    Best Local Similarity 100.0%; Pred. No. 1.7;
    Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 316 TCAGCAGAGCTGCTTCCCAAG 337
DB 450 TCAGCAGAGCTGCTTCCCAAG 429
RESULT 42
HUMPPRO/c
LOCUS
DEFINITION Homo sapiens (clone DN10mel) P protein mRNA, complete cds.
ACCESSION M99564
VERSION M99564.1 GI:190284
KEYWORDS P gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Metazoa;
Query Match 1..9%; Score 22; DB 9; Length 3070;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 316 TCAGCAGAGCTGCTTCCCAAG 337
DB 450 TCAGCAGAGCTGCTTCCCAAG 429
RESULT 43
BC012097/c
LOCUS
DEFINITION Homo sapiens oculocutaneous albinism II (pink-eye dilution homolog, mouse) mRNA (CDNA clone MGC:20070 IMAGE:4641135), complete cds.
ACCESSION BC012097
VERSION BC012097.1 GI:15082368
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3082)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3070)
Rinchik, E.M., Bultman, S.J., Horsthemke, B., Lee, S.T., Strunk, K.M.,
Spritz, R.A., Avidano, K.M., Jong, M.T. and Nicholls, R.D.
A gene for the mouse pink-eyed dilution locus and for human type II
oculocutaneous albinism
Nature 361 (6407), 72-76 (1993)
93133287
PUBMED
8421497
COMMENT
    Original source text: Homo sapiens (library: cDNA of R. Neve)
    fetal, and adult brain, and melanocyte cDNA to mRNA.
    Location/Qualifiers
        1..3070
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="brain, and melanocyte"
            /dev_stage="fetal, and adult"
            /tissue_lib="cDNA of R. Neve"
            1..52
            53..2569
            /codon_start=1
            /product="p protein"
            /protein_id="AAA36477.1"
            /db_xref="GI:190285"
            /translation="MHLEGRDGRYPGPAVELLQTSVPGLAELVAGKRLPRGAGG
            ADPSHCPRGAGQSSWAPAGQEFASFLLTKGRSHSLPQMSRSSRSDSCTFTPTLLR
            NSLQKSGRCIPVYHPEFTABESWSDADWERYLLSREVSGLSASASRKGDLDD
            SPHRLRLSKLRQVOMLVKGLFAFVLCISLFSLYDQGLKWLALSPLENYSVN
            LSSHVDSTLQVDLALGAVSGRPREHIVVELTODDGLSGRRRQOYTHNKTV
            YLNPRESHSVMSTFEVLTRETYSISIRASIQOQAVPLMAHQVLRGSVETQVTLA
            TAILAGVYALIFEIVHRTLAALGSLAALAVIGDRPSLTHVVVEWIDFETLALLF
            GMMLVAIFSETGFDYCAVKAYRLSRGRWAMIIMLCIIAAVLSAFLDNVTMLLFT
            PVTIRLCEVNLDPQVLIARVIFTNIGGAATAIGDPNVIIVSQELRKGDLDFAGF
            BAHMFIGICLVGLVCPFLRLLYWNRKLYNKEPSEIVELKHEIHVRLTAQRI SPASR
            BETAVRLLGLVLAHLRLRLHTFRQISQEDKNWETNIQELQKHRI SDGILLA
            KCLTVLGVIFEMFELNSFVPGHLDLGMIAIGAILLADIHDFEILHVEWATL
            LFPNALFVLMALAHLEHYVGEQTALLIKVPEQRRLIAILVIVVWSALASSLID
            NIPFTATMIPVLNLUSDPEVGLPAPPLMALAFAGCLGCGAAGTLLIGASANNVVCAGTAE
            QHGYSFMEFFELGFPFMMVSVCTGVMCYLLVAHVVVGVN"
            2570..3070
            3061..3066
            polyA_signal
ORIGIN
    Query Match 1..9%; Score 22; DB 9; Length 3070;
    Best Local Similarity 100.0%; Pred. No. 1.7;
    Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 316 TCAGCAGAGCTGCTTCCCAAG 337
DB 450 TCAGCAGAGCTGCTTCCCAAG 429
3'UTR
polyA_signal
ORIGIN
    Query Match 1..9%; Score 22; DB 9; Length 3070;
    Best Local Similarity 100.0%; Pred. No. 1.7;
    Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 316 TCAGCAGAGCTGCTTCCCAAG 337
DB 450 TCAGCAGAGCTGCTTCCCAAG 429
RESULT 43
BC012097/c
LOCUS
DEFINITION Homo sapiens oculocutaneous albinism II (pink-eye dilution homolog, mouse) mRNA (CDNA clone MGC:20070 IMAGE:4641135), complete cds.
ACCESSION BC012097
VERSION BC012097.1 GI:15082368
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3082)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3070)
Rinchik, E.M., Bultman, S.J., Horsthemke, B., Lee, S.T., Strunk, K.M.,
Spritz, R.A., Avidano, K.M., Jong, M.T. and Nicholls, R.D.
A gene for the mouse pink-eyed dilution locus and for human type II
oculocutaneous albinism
Nature 361 (6407), 72-76 (1993)
93133287
PUBMED
8421497
COMMENT
    Original source text: Homo sapiens (library: cDNA of R. Neve)
    fetal, and adult brain, and melanocyte cDNA to mRNA.
    Location/Qualifiers
        1..3070
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="brain, and melanocyte"
            /dev_stage="fetal, and adult"
            /tissue_lib="cDNA of R. Neve"
            1..52
            53..2569
            /codon_start=1
            /product="p protein"
            /protein_id="AAA36477.1"
            /db_xref="GI:190285"
            /translation="MHLEGRDGRYPGPAVELLQTSVPGLAELVAGKRLPRGAGG
            ADPSHCPRGAGQSSWAPAGQEFASFLLTKGRSHSLPQMSRSSRSDSCTFTPTLLR
            NSLQKSGRCIPVYHPEFTABESWSDADWERYLLSREVSGLSASASRKGDLDD
            SPHRLRLSKLRQVOMLVKGLFAFVLCISLFSLYDQGLKWLALSPLENYSVN
            LSSHVDSTLQVDLALGAVSGRPREHIVVELTODDGLSGRRRQOYTHNKTV
            YLNPRESHSVMSTFEVLTRETYSISIRASIQOQAVPLMAHQVLRGSVETQVTLA
            TAILAGVYALIFEIVHRTLAALGSLAALAVIGDRPSLTHVVVEWIDFETLALLF
            GMMLVAIFSETGFDYCAVKAYRLSRGRWAMIIMLCIIAAVLSAFLDNVTMLLFT
            PVTIRLCEVNLDPQVLIARVIFTNIGGAATAIGDPNVIIVSQELRKGDLDFAGF
            BAHMFIGICLVGLVCPFLRLLYWNRKLYNKEPSEIVELKHEIHVRLTAQRI SPASR
            BETAVRLLGLVLAHLRLRLHTFRQISQEDKNWETNIQELQKHRI SDGILLA
            KCLTVLGVIFEMFELNSFVPGHLDLGMIAIGAILLADIHDFEILHVEWATL
            LFPNALFVLMALAHLEHYVGEQTALLIKVPEQRRLIAILVIVVWSALASSLID
            NIPFTATMIPVLNLUSDPEVGLPAPPLMALAFAGCLGCGAAGTLLIGASANNVVCAGTAE
            QHGYSFMEFFELGFPFMMVSVCTGVMCYLLVAHVVVGVN"
            2570..3070
            3061..3066
            polyA_signal
ORIGIN
    Query Match 1..9%; Score 22; DB 9; Length 3070;
    Best Local Similarity 100.0%; Pred. No. 1.7;
    Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 316 TCAGCAGAGCTGCTTCCCAAG 337
DB 450 TCAGCAGAGCTGCTTCCCAAG 429

```

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Faney, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3082)

Strausberg, R.

Direct Submission

Submitted (02-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhru, Parvaneh Saedi, Jacqueline Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 29 Row: 1 Column: 19

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557810.

Location/Qualifiers

1. .3082

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:20070 IMAGE:4641135"

/tissue_type="skin, melanotic melanoma."

/clone_lib="NIH_MGC_20"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .3082

/gene="OCA2"

/note="synonyms: D15S12, PED, BOCA, P"

/db_xref="LOCUSID:4948"

/db_xref="MIM:203200"

11. .2555

/codon_start=1

/product="OCA2 protein"

/protein_id="AAH12097.1"

/db_xref="GI:15082369"

/db_xref="LOCUSID:4948"

/translation="MLEHGRDRRRPGAPAVELLQTSVPSGLAEIVACKRRLPRGAGG ADFHSCPGAGQSGWAPAGQFASFLTKGRSHSLPQMSSRSKSDSCFTENPLLR NSLOEKGSCIPVYHPEFTABESWEDSADWERYLLISRGVSGLSASAKGDLDD SPHRLRLSKARCVQWLMGLFAPVLCISFLYFDQGLWOLLALSPEKLVSVN LSHVDSTLQVDLACALVAGSPRGRREHIVELTQADLGSRRPQQVTHNVTY YLNPRESSEHSVMSRTFVLTRETIVTSISIRASLQQTQAVPPLLMAHGYLRGVSQTVTIA

TAILAGVVALIIFERPSTHVVWIDFETLALLFGMMILVAIFSETGFFDYCAVKAYR
LSRGVAMIMLCIAAVLSAFLDNVTMLLFTPTVIRLCEVLNLPQVLIAEVIF
TNIGGAATAIGDPNPVIVSNQELRKMGLDPAGFTAHMFICGLVLLVCPFLLLLYW
NKLXNKPSEIVELKHEIHWRLTAQRISPASREEVAVRLLGLKVLAEHLIARL
HTFHRQISOEDKNQWETIOEQKHRSIDGILLAKCLTVLGVFMFPLNSFVGIHL
DTWIAIGAILADILADHDPEIFILHREWATLFFAALFVLMELALHLLIYVGE
OTALLIKWVPEQRLLIAAIVLVVWSALASLIDNPFTATMIVLNLNLSHDPVGLP
APLMVALAFGACIGGCTGTCASANVVVCAIGAEQHGYSFSEFFRLGFFPMVVSCT
VGMCVLLVAHVVVVGN"

1107. .1580

/note="ArSB; Region: Na+/H+ antiporter NhaD and related
arsenite permeases [inorganic ion transport and
metabolism]"

/db_xref="CDD:COG1055"

1857. .2525

/note="ArSB; Region: Na+/H+ antiporter NhaD and related
arsenite permeases [inorganic ion transport and
metabolism]"

/db_xref="CDD:COG1055"

misc_feature

misc_feature

ORIGIN

Query Match 1.9%; Score 22; DB 9; Length 3082;
Best Local Similarity 100.0%; Freq. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 TCAGCAGAGCTGTCTTCCCAAG 337
|||||

Db 508 TCAGCAGAGCTGTCTTCCCAAG 487
|||||

RESULT 44

AC135349

LOCUS

DEFINITION

AC135349 Homo sapiens chromosome 17 clone RP13-910G3 map 17, LOW-PASS
SEQUENCE SAMPLING.

AC135349

ACCESSION

VERSION

HTG; HTGS_PHASE0.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 67718)

Birren, B., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 17, clone RP13-910G3

Unpublished

REFERENCE

2 (bases 1 to 67718)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Farro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,
Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norb, C.,
Norman, C. H., O'Connor, F., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Schojanovic, N., Talamas, J., Testaye, S., Theodore, J., Totham, K.,
Travers, M., Vasiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainov, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (12-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:
Snit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

TITLE

JOURNAL

COMMENT

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L28376
Center clone name: 910_G_3

* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
733: contig of 733 bp in length
734 833: gap of 100 bp
834 1567: contig of 734 bp in length
1568 1667: gap of 100 bp
1668 2404: contig of 737 bp in length
2405 2504: gap of 100 bp
2505 3211: contig of 707 bp in length
3212 3311: gap of 100 bp
3312 4062: contig of 751 bp in length
4063 4162: gap of 100 bp
4163 4895: contig of 733 bp in length
4896 4995: gap of 100 bp
4996 5738: contig of 743 bp in length
5739 5838: gap of 100 bp
5839 6579: contig of 741 bp in length
6580 6679: gap of 100 bp
6680 7409: contig of 730 bp in length
7410 7509: gap of 100 bp
7510 8263: contig of 754 bp in length
8264 8363: gap of 100 bp
8364 9100: contig of 737 bp in length
9101 9200: gap of 100 bp
9201 9933: contig of 733 bp in length
9934 10033: gap of 100 bp
10034 10769: contig of 736 bp in length
10770 10869: gap of 100 bp
10870 11590: contig of 721 bp in length
11591 11690: gap of 100 bp
11691 12428: contig of 738 bp in length
12429 12528: gap of 100 bp
12529 13271: contig of 743 bp in length
13272 13371: gap of 100 bp
13372 14106: contig of 735 bp in length
14107 14206: gap of 100 bp
14207 14949: contig of 743 bp in length
14950 15049: gap of 100 bp
15050 15794: contig of 745 bp in length
15795 15894: gap of 100 bp
15895 16553: contig of 759 bp in length
16554 17501: gap of 100 bp
17502 17501: contig of 748 bp in length
17503 18336: contig of 735 bp in length
18337 18436: gap of 100 bp
18437 19162: contig of 726 bp in length
19163 19262: gap of 100 bp
19263 20007: contig of 745 bp in length
20008 20107: gap of 100 bp
20108 20855: contig of 748 bp in length
20856 20955: gap of 100 bp
20956 21693: contig of 738 bp in length
21694 21793: gap of 100 bp
21794 22525: contig of 732 bp in length
22526 22625: gap of 100 bp
22626 23376: contig of 751 bp in length

23377 23476: gap of 100 bp
23477 24217: contig of 741 bp in length
24218 24317: gap of 100 bp
24318 25044: contig of 727 bp in length
25045 25144: gap of 100 bp
25145 25861: contig of 717 bp in length
25862 25961: gap of 100 bp
25962 26702: contig of 741 bp in length
26703 26802: gap of 100 bp
26803 27515: contig of 713 bp in length
27516 27615: gap of 100 bp
27616 28359: contig of 744 bp in length
28360 28459: gap of 100 bp
28460 29208: contig of 749 bp in length
29209 29308: gap of 100 bp
29310 30046: contig of 738 bp in length
30047 30146: gap of 100 bp
30147 30894: contig of 748 bp in length
30895 30994: gap of 100 bp
30996 31742: contig of 748 bp in length
31743 31842: gap of 100 bp
31843 32599: contig of 757 bp in length
32600 32699: gap of 100 bp
32700 33426: contig of 727 bp in length
33427 33526: gap of 100 bp
33527 34267: contig of 741 bp in length
34268 34367: gap of 100 bp
34368 35114: contig of 747 bp in length
35115 35214: gap of 100 bp
35215 35953: contig of 739 bp in length
35954 36053: gap of 100 bp
36054 36791: contig of 738 bp in length
36792 36891: gap of 100 bp
36892 37640: contig of 749 bp in length
37641 37740: gap of 100 bp
37741 38487: contig of 747 bp in length
38488 38587: gap of 100 bp
38588 39322: contig of 735 bp in length
39323 39422: gap of 100 bp
39423 40159: contig of 737 bp in length
40160 40259: gap of 100 bp
40260 41008: contig of 749 bp in length
41009 41108: gap of 100 bp
41109 41845: contig of 737 bp in length
41846 41945: gap of 100 bp
41946 42677: contig of 732 bp in length
42678 42777: gap of 100 bp
42778 43516: contig of 739 bp in length
43517 43616: gap of 100 bp
43617 44349: contig of 733 bp in length
44350 44449: gap of 100 bp
44450 45183: contig of 734 bp in length
45184 45283: gap of 100 bp
45284 46001: contig of 718 bp in length
46002 46101: gap of 100 bp
46102 46845: contig of 744 bp in length
46846 46945: gap of 100 bp
46946 47687: contig of 742 bp in length
47688 47787: gap of 100 bp
47788 48522: contig of 735 bp in length
48523 48622: gap of 100 bp
48623 49363: contig of 741 bp in length
49364 49463: gap of 100 bp
49464 50197: contig of 734 bp in length
50198 50297: gap of 100 bp
50298 51008: contig of 711 bp in length
51009 51108: gap of 100 bp
51109 51841: contig of 733 bp in length
51842 51941: gap of 100 bp
51942 52676: contig of 735 bp in length
52677 52776: gap of 100 bp
52777 53510: contig of 734 bp in length
53511 53610: gap of 100 bp

* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 733: contig of 733 bp in length
734 833: gap of 100 bp
834 1567: contig of 734 bp in length
1568 1667: gap of 100 bp
1668 2404: contig of 737 bp in length
2405 2504: gap of 100 bp
2505 3211: contig of 707 bp in length
3212 3311: gap of 100 bp
3312 4062: contig of 751 bp in length
4063 4162: gap of 100 bp
4163 4895: contig of 733 bp in length
4896 4995: gap of 100 bp
4996 5738: contig of 743 bp in length
5739 5838: gap of 100 bp
5839 6579: contig of 741 bp in length
6580 6679: gap of 100 bp
6680 7409: contig of 730 bp in length
7410 7509: gap of 100 bp
7510 8263: contig of 754 bp in length
8264 8363: gap of 100 bp
8364 9100: contig of 737 bp in length
9101 9200: gap of 100 bp
9201 9934: contig of 733 bp in length
9934 10033: gap of 100 bp
10034 10769: contig of 736 bp in length
10770 10869: gap of 100 bp
10870 11590: contig of 721 bp in length
11591 11690: gap of 100 bp
11691 12428: contig of 738 bp in length
12429 13271: contig of 743 bp in length
13272 13371: gap of 100 bp
13372 14106: contig of 735 bp in length
14107 14206: gap of 100 bp
14207 14949: contig of 743 bp in length
14950 15049: gap of 100 bp
15050 15794: contig of 745 bp in length
15795 15894: gap of 100 bp
15895 16653: contig of 759 bp in length
16654 16753: gap of 100 bp
16754 17501: contig of 748 bp in length
17502 18336: contig of 735 bp in length
18337 18436: gap of 100 bp
18437 19162: contig of 726 bp in length
19163 19263: gap of 100 bp
19264 20007: contig of 745 bp in length
20008 20107: gap of 100 bp
20108 20855: contig of 748 bp in length
20856 21693: contig of 738 bp in length
21694 21793: gap of 100 bp
21794 22525: contig of 732 bp in length
22526 22626: gap of 100 bp
22627 23376: contig of 751 bp in length
23377 23476: gap of 100 bp
23477 24217: contig of 741 bp in length
24218 24317: gap of 100 bp
24318 25044: contig of 727 bp in length
25045 25144: gap of 100 bp
25145 25861: contig of 717 bp in length
25862 25961: gap of 100 bp

Query Match 1.9%; Score 22; DB 2; Length 67718;
Best Local Similarity 100.0%; Pred.No. 1.2; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; TCCAGAGAGCTCTTCCAG 337
36997 TCCAGAGAGCTCTTCCAG 37018

AC135349 67718 bp DNA linear HTG 12-OCT-2002
Homo sapiens chromosome 17 clone RP13-910G3 map 17, LOW-PASS
SEQUENCE SAMPLING.
AC135349
AC135349.1 GI:23915532
HTG; HTGS PHASE0.
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 67718)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP13-910G3
Unpublished
2 (bases 1 to 67718)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagge,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
Mathews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

Submitted (12-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RN/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L28376
Center clone name: 910_G_3

TITLE
JOURNAL
COMMENT

* 25962 26702: contig of 741 bp in length
* 26703 26802: gap of 100 bp
* 27515: contig of 713 bp in length
* 27615: gap of 100 bp
* 27616 28359: contig of 744 bp in length
* 28359: contig of 744 bp in length
* 28459: gap of 100 bp
* 28460 29208: contig of 749 bp in length
* 29209 29308: gap of 100 bp
* 29309 30046: contig of 738 bp in length
* 30047 30146: gap of 100 bp
* 30147 30894: contig of 748 bp in length
* 30895 30994: gap of 100 bp
* 30995 31742: contig of 748 bp in length
* 31743 32599: contig of 757 bp in length
* 31843 32599: contig of 757 bp in length
* 32600 32899: gap of 100 bp
* 32700 33426: contig of 727 bp in length
* 33427 33526: gap of 100 bp
* 33527 34267: contig of 741 bp in length
* 34268 34367: gap of 100 bp
* 34368 35114: contig of 747 bp in length
* 35115 35214: gap of 100 bp
* 35215 35953: contig of 739 bp in length
* 35954 36053: gap of 100 bp
* 36054 36791: contig of 738 bp in length
* 36792 36891: gap of 100 bp
* 36892 37640: contig of 749 bp in length
* 37641 37740: gap of 100 bp
* 37741 38487: contig of 747 bp in length
* 38488 38587: gap of 100 bp
* 38588 39322: contig of 735 bp in length
* 39323 39422: gap of 100 bp
* 39423 40159: contig of 737 bp in length
* 40160 40259: gap of 100 bp
* 40260 41008: contig of 749 bp in length
* 41009 41108: gap of 100 bp
* 41109 41845: contig of 737 bp in length
* 41846 42677: contig of 732 bp in length
* 42678 42777: gap of 100 bp
* 42778 43516: contig of 739 bp in length
* 43517 43616: gap of 100 bp
* 43617 43499: contig of 733 bp in length
* 43499 44449: gap of 100 bp
* 44450 45183: contig of 734 bp in length
* 45184 45283: gap of 100 bp
* 45284 46001: contig of 718 bp in length
* 46002 46101: gap of 100 bp
* 46102 46845: contig of 744 bp in length
* 46846 46946: gap of 100 bp
* 46946 47687: contig of 742 bp in length
* 47688 47787: gap of 100 bp
* 47788 48522: contig of 735 bp in length
* 48523 48622: gap of 100 bp
* 48623 49364: contig of 741 bp in length
* 49364 49464: gap of 100 bp
* 49464 50197: contig of 734 bp in length
* 50198 50297: gap of 100 bp
* 50298 51008: contig of 711 bp in length
* 51009 51109: gap of 100 bp
* 51109 51841: contig of 733 bp in length
* 51842 51941: gap of 100 bp
* 51942 52676: contig of 735 bp in length
* 52677 52777: gap of 100 bp
* 52777 53510: contig of 734 bp in length
* 53511 53610: gap of 100 bp
* 53611 54348: contig of 738 bp in length
* 54349 54448: gap of 100 bp
* 54449 55191: contig of 743 bp in length
* 55192 55291: gap of 100 bp
* 55292 56040: contig of 749 bp in length
* 56041 56140: gap of 100 bp
* 56141 56867: contig of 727 bp in length

* 56868 56967: gap of 100 bp
* 56968 57703: contig of 736 bp in length
* 57704 57803: gap of 100 bp
* 57804 58557: contig of 754 bp in length
* 58558 58657: gap of 100 bp
* 58658 59406: contig of 749 bp in length

Query Match 1.9%; Score 22; DB 2; Length 67718;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 TCAGCAGAGCTGTCTTCCCAAG 337
|||||
Db 61472 TCAGCAGAGCTGTCTTCCCAAG 61451

RESULT 46
AC135345

LOCUS AC135345 155749 bp DNA linear HTG 12-NOV-2002
DEFINITION Homo sapiens chromosome 15 clone RP13-635J3 map 15, WORKING DRAFT
SEQUENCE, 9 unordered pieces.

AC135345
AC135345.2 GI:24899582
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 155749)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 15, clone RP13-635J3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 155749)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (12-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 155749)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

COMMENT
Direct Submission
Submitted (12-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 12, 2002 this sequence version replaced g1:23915528.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WISR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: 128371
Center clone name: 635 J.3
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 153532 bases at least Q40
Consensus quality: 154228 bases at least Q30
Consensus quality: 154617 bases at least Q20
Insert size: 156000; agarose-fp
Insert size: 154949; sum-of-contigs
Quality coverage: 16.6 in Q20 bases; agarose-fp
Quality coverage: 16.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 17950: contig of 17950 bp in length
* 17951: 18050: gap of 100 bp
* 18051: 18833: contig of 783 bp in length
* 18833: 18934: gap of 100 bp
* 18934: 20382: contig of 1449 bp in length
* 20382: 20483: gap of 100 bp
* 20483: 22317: contig of 1835 bp in length
* 22317: 22417: gap of 100 bp
* 22417: 23943: contig of 1526 bp in length
* 23943: 26418: gap of 100 bp
* 26418: 26518: gap of 100 bp
* 26518: 53742: contig of 27224 bp in length
* 53742: 53842: gap of 100 bp
* 53842: 97620: contig of 43778 bp in length
* 97620: 97720: gap of 100 bp
* 97720: 155749: contig of 58029 bp in length.

FEATURES
source
Location/Qualifiers
1. .155749
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP13-635J3"
/clone_lib="RPC1-13 Human Female BAC"

misc_feature
1. .17950
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
18051. .18833
/note="assembly_fragment"
18934. .20382
/note="assembly_fragment"
20483. .22317
/note="assembly_fragment"
22418. .23943
/note="assembly_fragment"

misc_feature
18051. .18833
/note="assembly_fragment"
18934. .20382
/note="assembly_fragment"
20483. .22317
/note="assembly_fragment"
22418. .23943
/note="assembly_fragment"

misc_feature
24044. .26418
/note="assembly_fragment"
misc_feature
26519. .53742
/note="assembly_fragment"
misc_feature
53843. .97620
/note="assembly_fragment"
misc_feature
97721. .155749
/note="assembly_fragment"

ORIGIN
Query Match 1.9%; Score 22; DB 2; Length 155749;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 316 TCAGCAGAGCTGCTTCCCAAG 337
Db 114280 TCAGCAGAGCTGCTTCCCAAG 114301

RESULT 47
AC124090 160842 bp DNA linear HTG 27-MAR-2003
LOCUS Homo sapiens chromosome 15 clone RP11-22H23 map 15, 3 unordered
DEFINITION AC124090
ACCESSION AC124090
VERSION AC124090.10 GI:27436826
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 160842)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 15, clone RP11-22H23
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 160842)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., LaRocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N.,
Mathews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (09-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 160842)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,
Mathews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,

* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 929: contig of 929 bp in length

1
 930 1029: gap of 100 bp
 1030 1726: contig of 697 bp in length
 1727 1826: gap of 100 bp
 1827 2511: contig of 685 bp in length
 2512 3271: contig of 660 bp in length
 3272 3371: gap of 100 bp
 3372 4067: contig of 696 bp in length
 4068 4167: gap of 100 bp
 4168 4786: contig of 619 bp in length
 4787 4886: gap of 100 bp
 4887 5512: contig of 626 bp in length
 5513 6588: contig of 976 bp in length
 6589 7557: contig of 1059 bp in length
 7558 7857: gap of 100 bp
 7859 8574: contig of 717 bp in length
 8575 8674: gap of 100 bp
 8676 9772: contig of 1098 bp in length
 9773 9872: gap of 100 bp
 9873 10552: contig of 680 bp in length
 10553 10852: gap of 100 bp
 10854 11392: contig of 740 bp in length
 11393 11432: gap of 100 bp
 11433 12728: contig of 1236 bp in length
 12729 12828: gap of 100 bp
 12829 13739: contig of 911 bp in length
 13740 13839: gap of 100 bp
 13840 14857: contig of 1018 bp in length
 14858 14957: gap of 100 bp
 14958 15734: contig of 777 bp in length
 15735 15834: gap of 100 bp
 15835 17247: contig of 1413 bp in length
 17248 17347: gap of 100 bp
 17348 19084: contig of 1737 bp in length
 19085 19184: gap of 100 bp
 19185 20354: contig of 1070 bp in length
 20355 20354: gap of 100 bp
 20356 22025: contig of 1671 bp in length
 22026 22125: gap of 100 bp
 22126 23667: contig of 1542 bp in length
 23668 25563: contig of 1756 bp in length
 25564 25663: gap of 100 bp
 25664 27849: contig of 2186 bp in length
 27850 27949: gap of 100 bp
 27950 28777: contig of 728 bp in length
 28778 30534: contig of 1757 bp in length
 30535 30634: gap of 100 bp
 30635 33577: contig of 2943 bp in length
 33578 33677: gap of 100 bp
 33678 34367: contig of 690 bp in length
 34368 36296: contig of 1829 bp in length
 36297 36396: gap of 100 bp
 36397 38987: contig of 2591 bp in length
 38988 39087: gap of 100 bp
 39088 41953: contig of 2866 bp in length
 41954 42053: gap of 100 bp
 42054 43622: contig of 1569 bp in length
 43623 43722: gap of 100 bp
 43723 46156: contig of 2434 bp in length
 46157 46256: gap of 100 bp

* 46257 49812: contig of 3556 bp in length
 * 49813 49912: gap of 100 bp
 * 49913 51975: contig of 2063 bp in length
 * 51976 52075: gap of 100 bp
 * 52076 56347: contig of 4272 bp in length
 * 56348 56447: gap of 100 bp
 * 56448 59548: contig of 3101 bp in length
 * 59549 59648: gap of 100 bp
 * 59649 61707: contig of 2059 bp in length
 * 61708 61807: gap of 100 bp
 * 61808 63850: contig of 2043 bp in length
 * 63851 63950: gap of 100 bp
 * 63951 67497: contig of 3547 bp in length
 * 67498 67597: gap of 100 bp
 * 67598 72596: contig of 4999 bp in length
 * 72597 72696: gap of 100 bp
 * 72697 76959: contig of 4263 bp in length
 * 76960 77059: gap of 100 bp
 * 77060 81851: contig of 4792 bp in length
 * 81852 81951: gap of 100 bp
 * 81952 85817: contig of 3866 bp in length
 * 85818 85917: gap of 100 bp
 * 85918 98207: contig of 12290 bp in length
 * 98208 98307: gap of 100 bp
 * 98308 109569: contig of 11262 bp in length
 * 109570 109669: gap of 100 bp
 * 109670 116612: contig of 6943 bp in length
 * 116613 116712: gap of 100 bp
 * 116713 124456: contig of 7744 bp in length
 * 124457 124556: gap of 100 bp
 * 124557 134424: contig of 9868 bp in length
 * 134425 134525: gap of 100 bp
 * 134526 147513: contig of 12989 bp in length
 * 147514 147614: gap of 100 bp
 * 147615 153189: contig of 5556 bp in length
 * 153190 153289: gap of 100 bp
 * 153290 158915: contig of 5646 bp in length
 * 158916 159015: gap of 100 bp
 * 159016 162195: contig of 3180 bp in length

Query Match 1.9%; Score 22; DB 2; Length 177613;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 316 TCAGCAGAGCTGCTTCCCAAG 337

Db 29851 TCAGCAGAGCTGCTTCCCAAG 29872

RESULT 49

AC017046/c

LOCUS

DEFINITION Homo sapiens chromosome 15 clone RP11-322N14, WORKING DRAFT

ACCESSION AC017046

VERSION AC017046.3 GI:7709950

KEYWORDS HTG; HTGS_PHADEL; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 180546)

Waterston, R.H.

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 180546)

Waterston, R.H.

Direct Submission

Submitted (09-DEC-1999) Genome Sequencing Center, Washington

University, School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On May 5, 2000 this sequence version replaced gi:7022553.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0322N14
----- Summary Statistics -----
Sequencing vector: M13; 83%
Chemistry: Dye-terminator Big Dye; 17% of reads
Chemistry: Dye-terminator Big Dye; 17% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 172064 bases at least Q40
Consensus quality: 175143 bases at least Q30
Consensus quality: 176933 bases at least Q20
Insert size: 171000; agarose-fp
Quality coverage: 179246; sum-of-contigs
Quality coverage: 4.50 in Q20 bases; agarose-fp
Quality coverage: 4.64 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1653: contig of 1653 bp in length
* 1654 1753: gap of unknown length
* 1754 5087: contig of 3334 bp in length
* 5088 5187: gap of unknown length
* 5188 8379: contig of 3191 bp in length
* 8379 8479 18069: contig of 9591 bp in length
* 8479 18070 18169: gap of unknown length
* 18070 26911: contig of 8742 bp in length
* 26911 27011: gap of unknown length
* 27011 37448: contig of 10437 bp in length
* 37448 37548: gap of unknown length
* 37548 48103: contig of 10555 bp in length
* 48103 48204 58987: contig of 10784 bp in length
* 48204 58987: gap of unknown length
* 58988 69005: contig of 9818 bp in length
* 69005 83140: gap of unknown length
* 83140 83240: contig of 14135 bp in length
* 83240 103712: contig of 20472 bp in length
* 103712 103812: gap of unknown length
* 103812 120089: contig of 16277 bp in length
* 120089 120189: gap of unknown length
* 120189 146597: contig of 28508 bp in length
* 146597 146797: gap of unknown length
* 146798 180546: contig of 33749 bp in length.
*
* Location/Qualifiers
* 1..180546
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="15"
* /clone="RP11-322N14"
* 1..1653
* /note="assembly_name:Contig14"
* 1754..5087
* /note="assembly_name:Contig15"
* 5188..8378
* /note="assembly_name:Contig17"
* 8479..18069
* /note="assembly_name:Contig18"
* 18170..26911
* /note="assembly_name:Contig19"
* 27012..37448

misc_feature
/note="assembly_name:Contig20"
37549..48103
/note="assembly_name:Contig21"
48204..58987
/note="assembly_name:Contig22"
59088..68905
/note="assembly_name:Contig23"
69006..83140
/note="assembly_name:Contig24"
83241..103712
/note="assembly_name:Contig25"
103813..120089
/note="assembly_name:Contig26"
clone end:SP6
vector side:right"
120190..146697
/note="assembly_name:Contig27"
146798..180546
/note="assembly_name:Contig28"

ORIGIN
Query Match 1.9% Score 22; DB 2; Length 180546;
Best Local Similarity 100.0%; Pred.No.1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 316 TCACGACAGCTGTCTTCCCAAG 337
|||||
Db 163941 TCACGACAGCTGTCTTCCCAAG 163920

RESULT 50
AC135329/c
LOCUS AC135329 Homo sapiens chromosome 15, clone RP11-1365A12, linear PRI 17-MAY-2003
DEFINITION AC135329
ACCESSION AC135329
VERSION AC135329.12 GI:30841056
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 188766)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
Miengwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainov,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (12-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 188766)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,

Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
McElrath, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 188766)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collimore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
McElrath, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (17-MAY-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 17, 2003 this sequence version replaced gi:29135660.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L28349
Center clone name: 1365_A_12

FEATURES
Source

Location/Qualifiers
1. .188766
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-1365A12"
/clone_lib="RP11-11 Human Male BAC segment 5"
complement(2032..2297)
/rpt_family="LIME1"
complement(2298..2589)
/rpt_family="AluJb"
complement(2590..4015)
/rpt_family="LIME1"
complement(4016..4313)
/rpt_family="AluY"
complement(4314..4459)

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/rpt_family="LIME1"
complement(471..5862)
/rpt_family="LIME2"
complement(5867..5990)
/rpt_family="LIME2"
7976..8173
/rpt_family="MER46C"
complement(8661..8787)
/rpt_family="LIME4a"
complement(8784..9282)
/rpt_family="LIME4a"
complement(9400..9786)
/rpt_family="LIME4a"
complement(9790..10026)
/rpt_family="LIME2"
complement(10030..10285)
/rpt_family="LIME4a"
complement(10286..10596)
/rpt_family="AluJb"
complement(10597..10772)
/rpt_family="LIME4a"
complement(13202..13497)
/rpt_family="AluY"
complement(13498..13774)
/rpt_family="LIME4a"
complement(13775..14059)
/rpt_family="AluX"
complement(14060..14217)
/rpt_family="LIME4a"
14363..14527
/rpt_family="MER3"
14833..14861
/rpt_family="AT-rich"
15225..21334
/rpt_family="LIME5"
complement(20817..20992)
/note="Single clone coverage"
complement(21414..21538)
/rpt_family="MIR"
complement(22434..22730)
/rpt_family="AluX"
complement(22742..23034)
/rpt_family="AluSp"
complement(24064..24543)
/rpt_family="LIME5"
24580..24622
/rpt_family="AT-rich"
complement(24626..24787)
/rpt_family="FRAM"
complement(28050..28355)
/rpt_family="AluJb"
complement(28481..28759)
/rpt_family="AluJb"
28846..29146
/rpt_family="L2"
29171..29229
/rpt_family="CT-rich"
complement(30497..30535)
/rpt_family="MER58"
32361..32666
/rpt_family="AluY"
34829..34865
/rpt_family="AT-rich"
complement(34867..35150)
/rpt_family="AluSg"
35153..35172
/rpt_family="TTTTTA"
35285..35566
/rpt_family="AluX"
complement(36013..36327)
/rpt_family="AluSg"
complement(36516..36924)
/rpt_family="MLTIF2"

```

repeat_region complement(36931..37055)
/rpt_family="Tigerr1"
repeat_region complement(37045..38771)
/rpt_family="Tigerr1"
repeat_region complement(39294..39488)
/rpt_family="MERS8A"
repeat_region 39897..39926
/rpt_family="AT_rich"
repeat_region 41134..41412
/rpt_family="ALUSx"
repeat_region 41415..42802
/rpt_family="L1M4"
repeat_region complement(42897..42942)

Query Match 1.9% Score 22; DB 9; Length 188766;
Best Local Similarity 100.0%; Pred.No.1.1; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 316 TCAGCAGAGCTGTCTCCCAAG 337
DB 148167 TCAGCAGAGCTGTCTCCCAAG 148146

RESULT 51
AC121324/c
LOCUS AC121324
DEFINITION Homo sapiens chromosome 15, clone RP11-959E3, complete sequence.
ACCESSION AC121324
VERSION AC121324.5 GI:22123153
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birren,B., Nussbaum,C. and Lander,E.
1 (bases 1 to 192087)
Homo sapiens chromosome 15, clone RP11-959E3
Unpublished
2 (bases 1 to 192087)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazates,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Minova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 192087)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Faro,S.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Minova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:22024591.
All repeats were identified using RepeatMasker:
Sait,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26781
Center clone name: 959_E_3
----- Location/Qualifiers
1. 192087
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosomes="15"
/map="15"
/clone="RP11-959E3"
/clone_lib="RP11-11 Human Male BAC"
/complement(1..79)
/rpt_family="AluS"
complement(546..726)
/rpt_family="L1ME3B"
complement(727..2129)
/rpt_family="L1MDa"
complement(2130..2367)
/rpt_family="MLT1B"
complement(2368..2672)
/rpt_family="AluSg"
complement(2673..2852)
/rpt_family="MLT1B"
complement(2853..3624)
/rpt_family="L1MDa"
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

```

```

Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Minova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 192087)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Faro,S.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Minova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:22024591.
All repeats were identified using RepeatMasker:
Sait,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26781
Center clone name: 959_E_3
----- Location/Qualifiers
1. 192087
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosomes="15"
/map="15"
/clone="RP11-959E3"
/clone_lib="RP11-11 Human Male BAC"
/complement(1..79)
/rpt_family="AluS"
complement(546..726)
/rpt_family="L1ME3B"
complement(727..2129)
/rpt_family="L1MDa"
complement(2130..2367)
/rpt_family="MLT1B"
complement(2368..2672)
/rpt_family="AluSg"
complement(2673..2852)
/rpt_family="MLT1B"
complement(2853..3624)
/rpt_family="L1MDa"
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

```


repeat_region complement(3679..3738)
 repeat_region /rpt_family="L1LMDa"
 repeat_region 4316..4424
 repeat_region /rpt_family="GA-rich"
 repeat_region 4673..4817
 repeat_region /rpt_family="MIR"
 repeat_region 6036..6359
 repeat_region /rpt_family="L1LME4a"
 repeat_region 6461..6585
 repeat_region /rpt_family="L1LME4a"
 repeat_region complement(7640..7931)
 repeat_region /rpt_family="AluJb"
 repeat_region complement(8122..8410)
 repeat_region /rpt_family="AluJo"
 repeat_region complement(9603..9921)
 repeat_region /rpt_family="AluSg"
 repeat_region 10657..10826
 repeat_region /rpt_family="MER20"
 repeat_region complement(11088..11649)
 repeat_region /rpt_family="LTR24C"
 repeat_region complement(12782..12952)
 repeat_region /rpt_family="MER3"
 repeat_region 13449..13749
 repeat_region /rpt_family="AluSx"
 repeat_region complement(15774..15840)
 repeat_region /rpt_family="MIR"
 repeat_region complement(16611..16708)
 repeat_region /rpt_family="L2"
 repeat_region 18445..18630
 repeat_region /rpt_family="(TG)n"
 repeat_region complement(20250..20524)
 repeat_region /rpt_family="L1LMBb"
 repeat_region complement(20749..20912)
 repeat_region /rpt_family="L1LMBb"
 repeat_region 23133..23217
 repeat_region /rpt_family="CT-rich"
 repeat_region complement(23377..23683)
 repeat_region /rpt_family="MER39"
 repeat_region 25604..25643
 repeat_region /rpt_family="AT-rich"
 repeat_region complement(25651..25926)
 repeat_region /rpt_family="AluJo"
 repeat_region 25930..26082
 repeat_region /rpt_family="MER369"
 repeat_region complement(26675..27034)
 repeat_region /rpt_family="L1LMB"
 repeat_region complement(28714..29067)
 repeat_region /rpt_family="MER68"
 repeat_region complement(29823..29865)
 repeat_region /rpt_family="LTR10C"
 repeat_region complement(30168..30452)
 repeat_region /rpt_family="L1LMB/D"
 repeat_region complement(30453..30492)
 repeat_region /rpt_family="Alu"
 repeat_region complement(30493..30791)
 repeat_region /rpt_family="AluSx"
 repeat_region complement(30792..30932)
 repeat_region /rpt_family="L1LMB/D"
 repeat_region complement(30933..31242)
 repeat_region /rpt_family="AluJo"
 repeat_region complement(31243..32412)
 repeat_region /rpt_family="L1LMB/D"
 repeat_region complement(32407..32696)
 repeat_region /rpt_family="L1LMB/D"
 repeat_region 32697..32740
 repeat_region /rpt_family="(TG)n"
 repeat_region complement(32741..33017)
 repeat_region /rpt_family="AluSx"
 repeat_region complement(33018..33099)
 repeat_region /rpt_family="L1LMB/D"
 repeat_region complement(33669..34262)
 repeat_region /rpt_family="LTR10C"
 repeat_region 34305..34389

repeat_region /rpt_family="L2"
 complement(34692..34976)
 repeat_region /rpt_family="AluSg"
 complement(36276..36426)
 repeat_region /rpt_family="MIR"
 36592..36961

Query Match 1.9%; Score 22; DB 9; Length 192087;
 Best Local Similarity 100.0%; Fred. No. 1.1;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 316 TCAGCAGAGCTGCTTCCCAAG 337
 |||||
 Db 14437 TCAGCAGAGCTGCTTCCCAAG 14416

RESULT 52
 AC135718 195397 bp DNA linear HTG 27-MAR-2003
 LOCUS Homo sapiens chromosome 15 clone RP11-1417P12 map 15, 3 ordered
 DEFINITION pieces.
 AC135718 GI:28301989
 VERSION HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_CANCELLED.
 KEYWORDS Homo sapiens
 SOURCE Homo sapiens
 ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 195397)
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 15, Clone RP11-1417P12
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 195397)
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferraira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,
 Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,
 Menga, V., Murphy, T., Navlor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zairoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 JOURNAL Submitted (21-OCT-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 195397)
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K.,
 Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferraira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Navlor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Kocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, R., Johnson, L., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, C., Kowals, C., Kratt, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuewa, L., Louisedg, H., Losado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathewine, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelemeh, O., Okwuonu, G., Olarunpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Sherty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steilme, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, M., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vava, V., Villanasa, D., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
Query Match	1.9%	Score 22;	DB 2;	Length 195397;		
Best Local Similarity	100.0%;	Pred. No. 1.1;				
Matches 22;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		

```
/clone lib="RPCI-11 Male BAC segment 5"
```

Query Match	1.9%;	Score 22;	DB 2;	Length 195397;
Best Local Similarity	100.0%;	Pred. No. 1.1;		
Matches 22;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	316	TCACGAGAGCTGCTTCCCAAG	337	
DB	37805	TCACGAGAGCTGCTTCCCAAG	37826	

RESULT	AC117033	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
53	AC117033	LOCUS	Rattus norvegicus clone CH230-259K17, WORKING DRAFT SEQUENCE.	AC117033	4	GI:25072994	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.	
							Rattus norvegicus (Norway rat)	
							Rattus norvegicus	
							Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	

REFERENCE
AUTHORS

1 (Bases 1 to 195895)
Murny, D., Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Anquiano, D.,
Anvabebechi, V., Avagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
ratus.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

```

----- Project Information
Center project name: GTO5
Center clone name: CH230-259K17
----- Summary Statistics
Assembly program: Phrap; version 0.930323
Consensus quality: 177612 bases at least Q40
Consensus quality: 179502 bases at least Q30
Consensus quality: 180779 bases at least Q20
Estimated insert size: 179376; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 1 contigs. Gaps between the contigs
  are represented as runs of N. The order of the pieces
  is believed to be correct as given, however the sizes
  of the gaps between them are based on estimates that have
  been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 195895: contig of 195895 bp in length.
  Location/Qualifiers
    1..195895
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-259K17"
    1..1399
      /note="wgs_end_extension"
      clone_end:77
      /note="wgs_end_extension"
      /note="wgs_end_extension"
      clone_end:77
      13904..15111
      /note="wgs_end_extension"
      clone_end:77
      complement(15863..16714)
      /note="clone_boundary"
      clone_end:77
      site:
      end_sequence:B2168558"
      complement(194830..195210)
      /note="clone_boundary"
      clone_end:Sp6
      site:
      end_sequence:B2168559"
-----
ORIGIN
Query Match 1.9%; Score 22; DB 2; Length 195895;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 CAAGAGGCGCTGGGCTTCCT 68
      |||||
Db 52645 CAAGAGGCGCTGGGCTTCCT 52666

RESULT 54
AC124091/c 196827 bp DNA linear PRI 08-JAN-2003
LOCUS Homo sapiens chromosome 15, clone RP11-640H21, complete sequence.
DEFINITION AC124091
ACCESSION AC124091
VERSION AC124091.16 GI:27545111
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 196827)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, clone RP11-640H21

```

```

JOURNAL REFERENCE
AUTHORS
Unpublished
2 (bases 1 to 196827)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collings,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LeRoque,K.,
Lamazares,B., Landers,T., Lehotzky,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,T., Nguyen,C.,
Nicoll,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL REFERENCE
AUTHORS
3 (bases 1 to 196827)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T.,
Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoll,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (03-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL REFERENCE
AUTHORS
4 (bases 1 to 196827)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T.,
Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoll,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-JAN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL REFERENCE
AUTHORS
On Jan 8, 2003 this sequence version replaced gi:26006646.
All repeats were identified using RepeatMasker.

```

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submission@genome.wi.mit.edu
 ----- Project Information
 Center project name: L26780
 Center clone name: 640_H_21

FEATURES source

Location/Qualifiers
 1. 196827

/organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-640H21"
 /clone_lib="RPC1-11 Human Male BAC"

repeat_region complement(2..1281)
 repeat_region /rpt_family="L1MDa"
 repeat_region complement(1282..1502)
 repeat_region /rpt_family="L1PA16"
 repeat_region complement(1503..1688)
 repeat_region /rpt_family="L1MDa"
 repeat_region 1689..2745
 repeat_region /rpt_family="MER11C"
 repeat_region complement(2746..3164)
 repeat_region /rpt_family="L1MDa"
 repeat_region complement(3165..3464)
 repeat_region /rpt_family="AluSc"
 repeat_region complement(3474..3774)
 repeat_region /rpt_family="AluJb"
 repeat_region complement(3775..5485)
 repeat_region /rpt_family="L1MDa"
 repeat_region complement(5486..6148)
 repeat_region /rpt_family="L2"
 repeat_region 6177..6758
 repeat_region /rpt_family="MLT1B"
 repeat_region complement(6923..7295)
 repeat_region /rpt_family="MLT1A2"
 repeat_region complement(7418..8152)
 repeat_region /rpt_family="MER45B"
 repeat_region complement(8759..9190)
 repeat_region /rpt_family="LTR12"
 repeat_region complement(10000..10322)
 repeat_region /rpt_family="L1PA16"
 repeat_region complement(10753..10872)
 repeat_region /rpt_family="L1ME"
 repeat_region complement(12428..12914)
 repeat_region /rpt_family="MLT1B"
 repeat_region 13088..13187
 repeat_region /rpt_family="LTR37A"
 repeat_region complement(13188..13443)
 repeat_region /rpt_family="AluSc"
 repeat_region 13444..13597
 repeat_region /rpt_family="LTR37A"
 repeat_region 13885..14186
 repeat_region /rpt_family="AluSp"
 repeat_region complement(14368..14677)
 repeat_region /rpt_family="AluJb8"
 repeat_region complement(15218..15492)
 repeat_region /rpt_family="L1ME1"
 repeat_region complement(15497..15894)
 repeat_region /rpt_family="L1ME1"
 repeat_region complement(15910..16389)
 repeat_region /rpt_family="L1ME1"
 repeat_region complement(16390..16545)
 repeat_region /rpt_family="AluSc"
 repeat_region complement(16546..16850)
 repeat_region /rpt_family="AluJ"
 repeat_region complement(16851..16986)

repeat_region /rpt_family="AluSc"
 complement(16987..17179)
 repeat_region /rpt_family="L1ME1"
 17213..17247
 repeat_region /rpt_family="TGn"
 complement(17249..17467)
 repeat_region /rpt_family="L1MD2"
 complement(17482..17816)
 repeat_region /rpt_family="L1PA16"
 complement(17817..18110)
 repeat_region /rpt_family="AluSc"
 complement(18111..18633)
 repeat_region /rpt_family="L1PA16"
 complement(18634..18935)
 repeat_region /rpt_family="AluJ"
 complement(19015..19157)
 repeat_region /rpt_family="AluJ"
 complement(19159..20057)
 repeat_region /rpt_family="L1"
 complement(20058..21800)
 repeat_region /rpt_family="SVA"
 21801..21887
 repeat_region /rpt_family="AGGGGGn"
 complement(21888..23014)
 repeat_region /rpt_family="L1"
 23012..23998
 repeat_region /rpt_family="L1P"
 complement(24005..24071)
 repeat_region /rpt_family="L1ME1"
 complement(24072..24938)
 repeat_region /rpt_family="MER4D"
 complement(24939..25279)
 repeat_region /rpt_family="L1ME1"
 25281..25574
 repeat_region /rpt_family="L1M4"
 25714..28095
 repeat_region /rpt_family="L1MB8"
 28168..29454
 repeat_region /rpt_family="L1MB8"
 complement(29456..29693)
 repeat_region /rpt_family="MLT1A1"
 complement(29704..30170)

Query Match 1.9%; Score 22; DB 9; Length 196827;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 316 TCAGCAGAGCTGCTTCCCAAG 337
 |||||
 Db 63255 TCAGCAGAGCTGCTTCCCAAG 63234

RESULT 55 AC098127

LOCUS AC098127 215835 bp DNA linear HTG 10-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-13416, WORKING DRAFT SEQUENCE.
 AC098127
 AC098127.7 GI:30522706
 HTG: HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 215835)
 Muzny, D. Marie, Metzker, M. Lee, Abranzon, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bardarona, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bliswalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.B., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huiyk, S., Hume, J., Idlebird, P., Jackson, A., Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowls, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulsegged, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mamoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokeneme, O., Okwundu, G., Olarpunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Primus, E., Pu, L.-L., Plopper, F., Poindexter, A., Popovic, D., Reeves, K., Regier, M.A., Reigh, R., Puaio, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sison, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorreller, R., Sosa, J., Steinkamp, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Unpublished
2 (bases 1 to 215835)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 215835)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE AUTHORS TITLE JOURNAL

On May 10, 2003 this sequence version replaced gi:22855453.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GGLV
Center clone name: CH230-13416
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 211950 bases at least Q40
Consensus quality: 212979 bases at least Q30
Consensus quality: 213496 bases at least Q20
Estimated insert size: 221121; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a working draft, sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 215835: contig of 215835 bp in length.

FEATURES
Location/Qualifiers
1..215835
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-13416"
1..1774
/note="wgs_contig"

ORIGIN
Query Match 1.9%; Score 22; DB 2; Length 215835;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 768 AGACCGAGGAGGCTTACTC 789
|||||
Db 81702 AGACCGAGGAGGCTTACTC 81723

RESULT 56
AC127197
LOCUS
DEFINITION
Rattus norvegicus clone CH230-232NS, WORKING DRAFT SEQUENCE, 2
unordered pieces.
AC127197 GI:24942133
VERSION
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 224266)
Muzny, D., Marie, D., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Bacca, F., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenship, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A., Chu, J., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Clee, A., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jarkpa, S., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, L., Loulseg, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, J., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmari, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 224266)
Worley, K.C.
Direct Submission
Submitted (14-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 224266)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23603939.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMSH
Center clone name: CH230-232N5
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 209253 bases at least Q40
Consensus quality: 211248 bases at least Q30
Consensus quality: 212355 bases at least Q20
Estimated insert size: 212371; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 222904: contig of 222904 bp in length
* 222905 223004: Gap of unknown length
* 223005 224266: contig of 1262 bp in length.

FEATURES
source
1..224266
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-232N5"
1..1631
/note="wgs_end_extension
clone_end:Sp6"
1682..3293
/note="wgs_end_extension
clone_end:Sp6"
5837..6363
/note="clone boundary
clone_end:Sp6
site:
end sequence:B2092052"
complement(213566..214446)
/note="clones_boundary
clone_end:T7
site:
end sequence:B2092052"
216090..217120
/note="wgs_end_extension
clone_end:T7"
221734..222904
/note="wgs_end_extension
clone_end:T7"

ORIGIN
Query Match 1.9%; Score 22; DB 2; Length 224266;
Best Local Similarity 100.0%; Pred.No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 47 CAAAGGGCCCTGGGCTTCCT 68
|||||
Db 178148 CAAAGGGCCCTGGGCTTCCT 178169
|||||

RESULT 57
AC128162/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-294H23, *** SEQUENCING IN PROGRESS
*** 3 unordered pieces.
AC128162 238907 bp DNA linear HTG 20-NOV-2002
Rattus norvegicus
AC128162 GI:25139364
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 238907)
Muzny, D., Marie, Metzker, M., Lee, Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Evans, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregorogor, B., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.B., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loulsegad, H., Lozador, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, N., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Primus, E., Pu, L.-L., Plopper, F., Polindexter, A., Popovic, D., Primm, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sittler, C.D., Smajd, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Hoit, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 238907)
Worley, K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 238907)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 20, 2002 this sequence version replaced gi:23908221. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZOG
Center clone name: CH230-294H23
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 218318 bases at least Q40
Consensus quality: 221565 bases at least Q30
Consensus quality: 223584 bases at least Q20
Estimated insert size: 226489; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 211726: contig of 211726 bp in length
* 211727 211826: gap of unknown length
* 211827 237283: contig of 25457 bp in length
* 237284 237383: gap of unknown length
* 237384 238907: contig of 1524 bp in length.
FEATURES
Location/Qualifiers
source
1..238907
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-294H23"
1..2339
misc_feature
/note="wgs_end_extension
clone_end:T7"
misc_feature
complement(3800..4593)
/note="clone_boundary
clone_end:T7"
site:
end_sequence:B2213875"
complement(5853..6712)
/note="clone boundary
clone_end:Sp6"
site:
end_sequence:B2213876"
211827..212946
/note="wgs_end_extension
clone_end:Sp6"
misc_feature
211827..212946
/note="wgs_end_extension
clone_end:Sp6"
ORIGIN
Query Match 1.9%; Score 22; DB 2; Length 238907;
Best Local Similarity 100.0%; Pred.No.1.1; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
QY 47 CAAAGGCGCCCTGGGCTTCCT 68
|||||
Db 155208 CAAAGGCGCCCTGGGCTTCCT 155187
RESULT 58
AX452882
LOCUS
DEFINITION
Sequence 3 from Patent WO0242457.
ACCESSION
AX452882
VERSION
AX452882.1 GI:21712521
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
artificial sequences.
REFERENCE
1

AUTHORS Chang,H., Yang,W.P., Wu,Y., Whitney,G.S., Perez-Villar,J.J. and Kanner,S.B.
TITLE Cloning and expression of human slap-2: a novel sh2/sh3 domain-containing human slap homologue having immune cell-specific expression
JOURNAL Patent: WO 0242457-A 3 30-MAY-2002; Bristol-Myers Squibb Co. (US)
FEATURES Location/Qualifiers
source 1..21
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="PY749 PCR PRIMER"

ORIGIN
Query Match 1.8%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred.No.11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 827 CGGATCAGACACTACAGGATC 847
|||||
Db 1 CGGATCAGACACTACAGGATC 21
|||||

RESULT 59
AC103328.1
WPCOMMENT
Sequence split into 4 fragments LOCUS AC103328 Accession AC103328
Fragment Name Begin End
AC103328_0 1 110000
AC103328_1 100001 210000
AC103328_2 200001 310000
AC103328_3 300001 355662
Continuation (2 of 4) of AC103328 from base 100001 (AC103328 Rattus norvegicus clone CH2)

Query Match 1.8%; Score 21; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred.No.4.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 477 CCATGGAGCAGAGAGAAGCA 497
|||||
Db 24296 CCATGGAGCAGAGAGAAGCA 24316
RESULT 60
AC137060/c
LOCUS AC137060 151416 bp DNA linear HTG 06-DEC-2002
DEFINITION Bos taurus clone RP42-222J11, WORKING DRAFT SEQUENCE, 3 ordered pieces.
ACCESSION AC137060
VERSION AC137060.2 GI:26080474
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (Bases 1 to 151416)
Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P., Lee-Lin,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J., Paquirigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished

TITLE
JOURNAL
REFERENCE 2 (Bases 1 to 151416)
AUTHORS Green,E.D.
TITLE Direct Submission

JOURNAL Submitted (15-NOV-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (Bases 1 to 151416)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT On Dec 6, 2002 this sequence version replaced gi:25013299.

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: edi
Center clone name: 222J11

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 151198 bases at least Q40
Consensus quality: 151215 bases at least Q30
Consensus quality: 151216 bases at least Q20
Insert size: 135000; agarose-fp
Insert size: 151216; sum-of-contigs
Quality coverage: 13.09x in Q20 bases; agarose-fp
Quality coverage: 11.69x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 41667: contig of 41667 bp in length
* 41668 41767: gap of unknown length
* 41768 107962: contig of 66195 bp in length
* 107963 108062: gap of unknown length
* 108063 151416: contig of 43354 bp in length.

FEATURES
Location/Qualifiers
source 1..151416
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-222J11"
/clone_lib="RP42"
1..78404
/note="clone overlaps with GenBank Accession Number AC137532 clone RP42-314C8 (center project name edh)"
misc_feature 1..41667
/note="assembly_fragment
clone_end:77
vector_side:left"
41768..107962
/note="assembly_fragment"
108063..151416
/note="assembly_fragment"

```
clone_end:SP6
vector_side:right"
```

```
Query Match      1.8%; Score 21; DB 2; Length 151416;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 840 ACAGGATCCACTGCCTTGACA 860
|||
Dp 42093 ACAGGATCCACTGCCTTGACA 42073

RESULT 61					
AC119240					
LOCUS	156129 bp	DNA	linear	HTG 10-MAR-2003	
DEFINITION	Mus musculus clone RP24-328K15,	WORKING DRAFT SEQUENCE,	6 unordered		
	pieces.				

TITLE
JOURNAL
Direct Submission
Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 156129)
PREFERENCE

AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Archachi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collimore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Poh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stofanovic, N., Stubbs, M.,

Talamas, J., Tsafaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkatraman, V.S., Viel, R., Vo, A.K., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. (2003) Direct Submission
Submitted (10-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 10, 2003 this sequence version replaced gi:28191472.
All repeats were identified using RepeatMasker:
Smith, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Genome Center

ORIGIN

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-23 BAC Library has been constructed by Kazutoyo Oseawa and Minako Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES

Source

1. .169618

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="12"

/map="12"

/clone="RP23-281J12"

/clone_lib="RPCI-23"

4619..4794

/rpt_family="B2"

4885..4994

/rpt_family="Alu"

5895..6022

/rpt_family="Alu"

6116..6292

/rpt_family="B2"

6316..6503

/rpt_family="B2"

complement(6426..6497)

/product="tRNA-Ser"

/note="Likely pseudogene (HMM SC=34.93 / Sec struct

SC=-10.52)"

6890..7060

/rpt_family="B4"

7548..7633

/rpt_family="Alu"

7649..7720

/rpt_family="ID"

7955..8010

/rpt_family="Alu"

8070..8225

/rpt_family="B4"

11852..11923

/rpt_family="ID"

11936..12140

/rpt_family="B2"

12778..13149

/rpt_family="RMR17C"

13872..13976

/rpt_family="Alu"

14138..14311

/rpt_family="B2"

16057..16148

/rpt_family="Alu"

18833..19801

/note="CpG island (GC=67.7, o/e=0.84, #CpGs=97)"

20719..20823

/rpt_family="Alu"

20761..20925

/rpt_family="B4"

22735..22929

/rpt_family="B2"

22965..23167

/rpt_family="B2"

23323..23512

/rpt_family="MER1_type"

23612..23904

repeat_region

repeat_region

Query Match 1.8%; Score 21; DB 2; Length 156129;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1068 TCCTGTTTCTGAAGCTGCCA 1088
|||||
Db 120261 TCCTGTTTCTGAAGCTGCCA 120281

RESULT 62
AC122811/c 169618 bp DNA linear ROD 08-NOV-2003
DEFINITION Mus musculus BAC clone RP23-281J12 from 12, complete sequence.

AC122811
ACCESSION AC122811.4 GI:23334945

VERSION HTG.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 169618)

Buatsi, D., Shah, N., Haakenson, W. and Creason, K.

The sequence of Mus musculus BAC clone RP23-281J12

Unpublished (2001)

2 (bases 1 to 169618)

Wilson, R.

Sequencing of Mus musculus

Unpublished (2001)

3 (bases 1 to 169618)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 169618)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 169618)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (27-SEP-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

6 (bases 1 to 169618)

Wilson, R.

Direct Submission

Submitted (08-NOV-2003) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Sep 27, 2002 this sequence version replaced gi:22475546.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BA0281J12

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping

clone sections once, or longer because we provide a small overlap

between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence

from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

```

/rpt_family="MaLR"
23871..23970
/rpt_family="ERVL"
23971..24158
/rpt_family="B2"
24159..24189
/rpt_family="ERVL"
24190..24262
/rpt_family="MaLR"
24312..24456
/rpt_family="B4"
24563..24663
/rpt_family="Alu"
26326..26514
/rpt_family="B2"
26518..26907
/rpt_family="L1"
26975..27049
/rpt_family="ID"
27139..27563
/rpt_family="MaLR"
27705..27890
/rpt_family="B2"
27891..28078
/rpt_family="B2"
28079..28099
/rpt_family="B2"
28172..28226
/rpt_family="Alu"
28231..28357
/rpt_family="B4"
28677..28731
/rpt_family="MIR"
29241..29561
/rpt_family="MaLR"
29644..29797
/rpt_family="Alu"
30597..30723
/rpt_family="B4"
30809..31076
/rpt_family="MaLR"
31318..31478
/rpt_family="MaLR"
32369..32763
/rpt_family="MaLR"
32814..32960
/rpt_family="Alu"
35710..35849
/rpt_family="MaLR"
41681..41830
/rpt_family="Alu"
41993..42094
/rpt_family="Alu"
42095..42128
/rpt_family="B4"
43995..44069
/rpt_family="ID"
45018..45171
/rpt_family="L1"
45375..45501
/rpt_family="Alu"
45992..46092
/rpt_family="U6"
46781..46972
/rpt_family="B2"
46997..47194

```

Query Match 1.8%; Score 21; DB 10; Length 169618;
 Best Local Similarity 100.0%; Pred.No. 4.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 477 CCATGGAAGCAGAGAGCA 497
 |||||

Db 82135 CCATGGAAGCAGAGAGCA 82115

RESULT 63
 AC144901/c

LOCUS 178293 bp DNA linear HTG 22-JUL-2003
 SUS scrofa clone RP44-138L19, WORKING DRAFT SEQUENCE.

DEFINITION AC144901

AC144901.3 GI:331112720

VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS SUS scrofa (pig)

SOURCE SUS scrofa

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 178293)

AUTHORS

Antoniellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Carlsberg,K., Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,O.L., Maduro,V.B., Margulies,E.H., Masiello,C., Maskari,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Reddix-Dugue,N., Schandier,K., Schueler,M.G., Shan,K., Sison,C., Stantropop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

UNPUBLISHED

2 (bases 1 to 178293)

Green,E.D.

Direct Submission

Submitted (24-MAY-2003) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 178293)

Green,E.D.

Direct Submission

Submitted (22-JUL-2003) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

On Jul 22, 2003 this sequence version replaced gi:32306561.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zoo@ngri.nih.gov

----- Project Information

Center project name: ecc

Center clone name: 138119

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 178287 bases at least Q40

Consensus quality: 178293 bases at least Q30

Consensus quality: 178293 bases at least Q20

Insert size: 167000; agarose-fp

Insert size: 178293; sum-of-contigs

Quality coverage: 11.38x in Q20 bases; agarose-fp

Quality coverage: 10.66x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 178293: contig of 178293 bp in length.

FEATURES

Source
 1..178293
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="RP44-138L19"
 /clone_lib="RP44"
 1..178293
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 clone_end:T7
 vector_side:right
 1..83454
 /note="clone overlaps with GenBank Accession Number
 AC145413 clone RP44-386P23 (center project name ecdb)"

ORIGIN

Query Match 1.8%; Score 21; DB 2; Length 178293;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 840 ACAGGATCCACTGCCTTGACA 860
 |||||
 Db 42418 ACAGGATCCACTGCCTTGACA 42398

RESULT 64
 AC137532/c
 LOCUS AC137532 181245 bp DNA linear HTG 18-DEC-2002
 DEFINITION Bos taurus clone RP42-314C8, WORKING DRAFT SEQUENCE, 3 ordered
 pieces.
 AC137532
 AC137532.2 GI:27228814
 VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.
 KEYWORDS Bos taurus (cow)
 SOURCE
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 181245)
 Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
 Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
 Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
 Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
 Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
 Margulies,E.H., Masello,C., Maskeri,B., McDowell,J.,
 Paquirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
 Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
 Stantropop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
 Wecherby,K.D., Wiggins,L., Young,A. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished

TITLE
 JOURNAL
 REFERENCE 2 (bases 1 to 181245)
 AUTHORS Green,E.D.
 TITLE Direct Submission
 JOURNAL Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 REFERENCE 3 (bases 1 to 181245)
 AUTHORS Green,E.D.
 TITLE Direct Submission

TITLE
 JOURNAL
 REFERENCE 18-DEC-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 COMMENT On Dec 18, 2002 genome version replaced gi:25167111.
 ----- Genome Center

Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: edh
 Center clone name: 314C08

The sequence data in this record represents an 'enhanced'
 version of a Phase 2 submission. Specifically, the indicated
 order and orientation of each sequence contig has been
 established using one or more of the following: read-pair
 data from individual subclones, overlaps with neighboring
 clones, alignment with available reference sequence (e.g.,
 human), and/or confirmation by PCR testing. In addition,
 the sequence assembly is based on at least 8X average
 coverage in Q20 bases and has been reviewed to rule out
 gross misassemblies, the low-quality ends of sequence
 contigs have been trimmed away, and each base is associated
 with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 180924 bases at least Q40
 Consensus quality: 181014 bases at least Q30
 Consensus quality: 181030 bases at least Q20
 Insert size: 155000; agarose-fp
 Insert size: 181045; sum-of-contigs
 Quality coverage: 12.75x in Q20 bases; agarose-fp
 Quality coverage: 10.91x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 144326: contig of 144326 bp in length
 * 144327 144426: gap of unknown length
 * 144427 146712: contig of 2286 bp in length
 * 146713 146812: gap of unknown length
 * 146813 181245: contig of 34433 bp in length.

FEATURES

Source
 1..181245
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-314C8"
 /clone_lib="RP42"
 1..144326
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 102736..181245
 /note="clone overlaps with GenBank Accession Number
 AC137060 clone RP42-222J11 (center project name edi)"
 144427..146712
 /note="assembly_fragment"
 146813..181245
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right

ORIGIN

Query Match 1.8%; Score 21; DB 2; Length 181245;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 840 ACAGGATCCACTGCCTTGACA 860

Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melgrim, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (10-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 10, 2003 this sequence version replaced gi:32452555.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L29072

Center clone name: 342_B.21

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 184598 bases at least Q40

Consensus quality: 185965 bases at least Q30

Consensus quality: 186687 bases at least Q20

Insert size: 185000; agarose-fp

Insert size: 187643; sum-of-contigs

Quality coverage: 8.9 in Q20 bases; agarose-fp

Quality coverage: 8.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 57438: contig of 57438 bp in length
* 57439 57538: gap of 100 bp
* 57539 58018: contig of 480 bp in length
* 58019 58118: gap of 100 bp
* 58119 58283: contig of 1165 bp in length
* 58284 59383: gap of 100 bp
* 59384 61298: contig of 1915 bp in length
* 61299 61399: gap of 100 bp
* 61399 62841: contig of 1443 bp in length
* 62842 62942: gap of 100 bp
* 62942 63166: contig of 2225 bp in length
* 63166 65267: gap of 100 bp
* 65267 67943: contig of 2676 bp in length
* 67943 68043: gap of 100 bp
* 68043 68460: contig of 418 bp in length
* 68461 68561: gap of 100 bp
* 68561 71225: contig of 8565 bp in length
* 71226 77225: gap of 100 bp
* 77226 100280: contig of 23055 bp in length
* 100281 100381: gap of 100 bp
* 100381 128542: contig of 28162 bp in length
* 128543 128643: gap of 100 bp
* 128643 185951: contig of 57309 bp in length
* 185952 186052: gap of 100 bp
* 186052 188843: contig of 2792 bp in length.

```

Location/Qualifiers

```

source
1. 188843
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="15"
/map="15"
/clone="RP23-342B21"
/clone_lib="RPC1-23 Female Mouse BAC"
1. 57438
/note="assembly_fragment"
vector_side:left
clone_end:SP6
57539-58018
/note="assembly_fragment"
58119-59283
/note="assembly_fragment"
59384-61298
/note="assembly_fragment"
61399-62841
/note="assembly_fragment"
62942-65166
/note="assembly_fragment"
65267-67942
/note="assembly_fragment"
68043-68460
/note="assembly_fragment"
68561-77125
/note="assembly_fragment"
77226-100280
/note="assembly_fragment"
100381-128542
/note="assembly_fragment"
128643-185951
/note="assembly_fragment"
186052-188843
/note="assembly_fragment"
clone_end:T7
vector_side:right"

misc_feature
1. 57438

misc_feature
57539-58018

misc_feature
58119-59283

misc_feature
59384-61298

misc_feature
61399-62841

misc_feature
62942-65166

misc_feature
65267-67942

misc_feature
68043-68460

misc_feature
68561-77125

misc_feature
77226-100280

misc_feature
100381-128542

misc_feature
128643-185951

misc_feature
186052-188843

vector_side:right"

```

ORIGIN

Query Match 1.8%; Score 21; DB 2; Length 188843;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 268 TCCAGAGGCGCCCAAGCC 288

Db 20951 TCCAGAGGCGCCCAAGCC 20971

RESULT 67

AC133116

LOCUS

DEFINITION AC133116 198893 bp DNA linear HTG 20-NOV-2002
 Rattus norvegicus clone CH230-499P20, WORKING DRAFT SEQUENCE, 2

AC133116

AC133116.3 GI:25138979

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 198893)

REFERENCE

AUTHORS

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buha, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacobs, L., Jiang, H., Johnson, B., Johnson, R., Joliviet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loresuhewa, L., Loulseghe, H., Lozado, R., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwunu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rokey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valdes, R., Vera, V., Villagana, D., Waldron, L., Walker, S., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 198893)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (07-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 198893)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:22771315.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
----- Project name: KCCE

Center clone name: CH230-499P20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 186018 bases at least Q40
Consensus quality: 188352 bases at least Q30
Consensus quality: 189734 bases at least Q20
Estimated insert size: 193933; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 84930: contig of 84830 bp in length
* 84931 84930: gap of unknown length
* 84931 198893: contig of 113963 bp in length.
----- Location/Qualifiers
1..198893
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-499P20"
3003..3641
/note="clone_boundary"
clone_end:77
site:
end sequence:B272062"
complement(80087..81000)
/note="clone_boundary"
clone_end:sp6
site:
end sequence:B272063"
84931..86604
/note="wgs_end_extension"
clone_end:sp6"

FEATURES
source
misc_feature
misc_feature
misc_feature
ORIGIN
Query Match 1.8%; Score 21; DB 2; Length 198893;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 477 CMTGGAGCAGAGAGAGCA 497
|||||
Db 41848 CCATGGAGCAGAGAGCA 41868

RESULT 68
AC145413/c
LOCUS
DEFINITION
AC145413
AC145413.2 GI:33413343
HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE
Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 206826)
AUTHORS
Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, A.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, X.,
Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,
Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
Hurle, B., Idol, J.R., Karlins, E., Kwong, P., Lalic, P., Lee-Lin, S.-Q.,
Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C.,
Maskeri, B., McDowell, J., Pagnirangan, C., Pearson, R., Porcnoy, M.E.,
Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Shah, K.,

Sison, C., Stantiprop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
 Vogt, J.L., Wetherby, K.D., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 206826)
 Green, E.D.
 Direct Submission
 Submitted (02-JUL-2003) NIH Intramural Sequencing Center, 8717
 Grovmont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 206826)
 Green, E.D.
 Direct Submission
 Submitted (02-AUG-2003) NIH Intramural Sequencing Center, 8717
 Grovmont Circle, Gaithersburg, MD 20877, USA
 On Aug 2, 2003 this sequence version replaced gi:32401577.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: scb
 Center clone name: 386F23

The sequence data in this record represents an 'enhanced'
 version of a Phase 2 submission. Specifically, the indicated
 order and orientation of each sequence contig has been
 established using one or more of the following: read-pair
 data from individual subclones, overlaps with neighboring
 clones, alignment with available reference sequence (e.g.,
 human), and/or confirmation by PCR testing. In addition,
 the sequence assembly is based on at least 8x average
 coverage in Q20 bases and has been reviewed to rule out
 gross misassemblies; the low-quality ends of sequence
 contigs have been trimmed away, and each base is associated
 with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 206259 bases at least Q40
 Consensus quality: 206416 bases at least Q30
 Consensus quality: 206502 bases at least Q20
 Insert size: 185000; agarose-fp
 Insert size: 206526; sum-of-contigs
 Quality coverage: 10.84x in Q20 bases; agarose-fp
 Quality coverage: 9.71x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 46921: contig of 46921 bp in length
 * 46922 47021: gap of unknown length
 * 47022 136987: contig of 89966 bp in length
 * 136988 137087: gap of unknown length
 * 137088 159322: contig of 22235 bp in length
 * 159323 159423: gap of unknown length
 * 159423 206826: contig of 47404 bp in length.

FEATURES

Location/Qualifiers
 1..206826
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="RP44-386P23"
 /clone_lib="RP44"
 1..46921
 /note="assembly_fragment"

misc_feature

clone_end:SP6
 vector_side:left
 47022..136987
 /note="assembly_fragment"
 124320..206826
 /note="clone overlaps with GenBank Accession Number
 AC144901 clone RP44-138L19 (center project name ecc)"
 137088..159322
 /note="assembly_fragment"
 159423..206826
 /note="assembly_fragment"
 clone_end:rv
 vector_side:right

ORIGIN

Query Match 1.8%; Score 21; DB 2; Length 206826;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 840 ACAGGATCCACTGCTTGACA 860
 Db 165790 ACAGGATCCACTGCTTGACA 165770

RESULT 69

AC098767/c

LOCUS

DEFINITION Rattus norvegicus clone CH230-154K15, WORKING DRAFT SEQUENCE, 7
 213182 bp DNA linear HTG 10-MAY-2003

ACCSSION

AC098767.5 GI:30521109

VERSION

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 213182)

Muzny, D.M., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyaiebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,

Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraut, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Loresu, H., Lou, X., Lu, X., Lu, X., Ma, J.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, K., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, T., Sitter, C.D., Smajda, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, X., Zhou, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiser, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 213182)
Worley, K.C.
Direct Submission
Submitted (01-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 213182)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23268900.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRRP
Center clone name: CH230-154KL15
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 200574 bases at least Q40
Consensus quality: 203581 bases at least Q30
Consensus quality: 205897 bases at least Q20
Estimated insert size: 211214; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 126909: contig of 126909 bp in length
* 126910 127009: gap of unknown length
* 127010 139859: contig of 12850 bp in length
* 139860 139959: gap of unknown length
* 139960 206171: contig of 66212 bp in length
* 206172 206271: gap of unknown length

* 206272 207428: contig of 1157 bp in length
* 207429 207528: gap of unknown length
* 207529 208536: contig of 1008 bp in length
* 208537 208636: gap of unknown length
* 208637 209959: contig of 1323 bp in length
* 209960 210059: gap of unknown length
* 210060 213182: contig of 3123 bp in length.

FEATURES
Location/Qualifiers
1..213182
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-154KL15"
1851..2512
/note="clone boundary
clone end:T7"
site:ECORI
end sequence:RWBM68TJB"
127010..128390
/note="wgs_contig"

misc_feature
1851..2512

misc_feature
127010..128390

ORIGIN
Query Match 1.8%; Score 21; DB 2; Length 213182;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 477 CCATGGAGCAGAGAGAGCA 497
|||||
Db 170769 CCATGGAGCAGAGAGCA 170749
|||||

RESULT 70
AC118839/c
LOCUS AC118839 222292 bp DNA linear HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-232222, WORKING DRAFT SEQUENCE, 3
unordered pieces.
ACCESSION AC118839
VERSION AC118839.4 GI:25137840
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 222292)
Muzny, D., Maric, M., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Dengson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhsu, L., Louised, H., Lozada, R.J., Lu, X., Lu, X., Lu, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,

Milosevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mucic, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quintero, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, B., Umanik, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE JOURNAL REFERENCE AUTHORS JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 222292)
Worley, K.C.

Submitted (21-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 222292)
Rat Genome Sequencing Consortium.

REFERENCE AUTHORS JOURNAL

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23908256.

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GVUI
Center clone name: CH230-232J22
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 214049 bases at least Q40
Consensus quality: 215716 bases at least Q30
Consensus quality: 216698 bases at least Q20
Estimated insert size: 219947; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

* be preserved.

* 1 120010: contig of 120010 bp in length
* 120011 120110: gap of unknown length
* 120111 220110: contig of 100000 bp in length
* 220111 220210: gap of unknown length
* 220211 222292: contig of 2082 bp in length.

FEATURES source

Location/Qualifiers
1..222292

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"

/clone="CH230-232J22"

2483..3345

/note="clone boundary"

clone_end:T7

site:

end sequence:RWBP59TJ"

complement(118194..119112)

/note="clone boundary"

clone_end:Sp6

end sequence:RWBP59TV"

120111..121353

/note="wgs end extension"

clone_end:Sp6"

ORIGIN

Query Match 1.8%; Score 21; DB 2; Length 222292;

Best Local Similarity 100.0%; Pred. No. 4.2;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 477 CCATGGAGCAGAGAGAGCA 497

Db 116546 CCATGGAGCAGAGAGAGCA 116526

RESULT 71

AC133973/c

LOCUS

DEFINITION

AC133973

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 247989)

Muzny, D., Maric, M., Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alshrocks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, I., Garza, M., Gregor, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Gunaratne, P., Havel, P., Hawes, A., Henderson, N., Hernandez, J., Harvey, Y., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hui, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

247989 bp DNA linear HTG 10-MAY-2003

Rattus norvegicus clone CH230-24P19, *** SEQUENCING IN PROGRESS

AC133973

GI:30522747

HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Rattus norvegicus

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 247989)

Muzny, D., Maric, M., Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alshrocks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, I., Garza, M., Gregor, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Gunaratne, P., Havel, P., Hawes, A., Henderson, N., Hernandez, J., Harvey, Y., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hui, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, J., Loulsged, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, S., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, M., Norris, S., Nwaakelen, O., Okwunnu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Popper, P., Poudexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherz, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajis, D., Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, I., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, S., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, J., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 247989)
Rat Genome Sequencing Consortium.
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 247989)
Rat Genome Sequencing Consortium.
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23264284.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSKU
Center clone name: CH230-24P19
----- Summary Statistics
Assembly program: Atlas 3.0
Consensus quality: 231477 bases at least Q40
Consensus quality: 235116 bases at least Q20
Consensus quality: 237974 bases at least Q20
Estimated insert size: 241770; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 247989: contig of 247989 bp in length.

FEATURES
Location/Qualifiers
source
1..247989
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-24P19"
1..1241
misc_feature
/note="wgs_contig"
misc_feature
246253..247989
/note="wgs_contig"
ORIGIN
Query Match 1.8%; Score 21; DB 2; Length 247989;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 586 TGAGGATGGAGCTGGTGAC 606
DB 13452 TGAGGATGGAGCTGGTGAC 13432

RESULT 72
AC130977/c
LOCUS
AC130977 249671 bp DNA linear HTG 10-MAY-2003
Rattus norvegicus clone CH230-117P2, WORKING DRAFT SEQUENCE, 2
unordered pieces.
AC130977 GI:30521555
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLPOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 249671)
Muzny, D.M., Metzker, M., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, X., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, J., Lu, X., Lu, Y., Lozano, R., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mapua, P., Martin, K., Martin, R., Martinez, S., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwankwelu, O., Okunribido, A., Olamunigbon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Plankko, C.,
Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
Williams, G., Willson, R., Wiczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 249671)
Rat Genome Sequencing Consortium.
Submitted (16-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 249671)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:25008794.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHLI
Center clone name: CH230-117P2
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 235617 bases at least Q40
Consensus quality: 235512 bases at least Q30
Consensus quality: 237119 bases at least Q20
Estimated insert size: 243815; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 2 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 248113: contig of 248113 bp in length

* 248114 248213: gap of unknown length
* 248214 249671: contig of 1458 bp in length.

FEATURES
source
1. 249671
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-117P2"
2039. 2742
/note="clone boundary
clone end:T7
site:ECORI
end_sequence: BH284352"
misc_feature
complement(246807..247531)
/note="clone boundary
clone end:Sp6
site:ECORI
end_sequence: BH284354"
ORIGIN

Query Match 1.8%; Score 21; DB 2; Length 249671;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1135 CTTCATCATCAGCTGAATGA 1155
|||||
Db 33266 CTTCATCATCAGCTGAATGA 33246
|||||

RESULT 73
AC102993/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-66D7, *** SEQUENCING IN PROGRESS ***;
7 unordered pieces.
AC102993
AC102993.6 GI:30521808
VERSION
HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 254581)
Muzny, D., Marle, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerrero, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, T., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Herrandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kow, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, X., London, P., Longacre, S., Lopez, J.,
Lorenshuwa, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
 Nwaokemele, O., Okunolu, G., Olampong, A., Pal, S., Pal, S.,
 Paternak, S., Paul, H., Perez, A., Perez, L., Prankoch, C.,
 Popper, F., Poudexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsheyn, A., Sisson, I., Sitter, C.D., Smajd, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, D.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, S., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczek, R., Wood, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiser, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 254581)
 Worley, K.C.

Direct Submission
 Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 254581)
 Rat Genome Sequencing Consortium.

Direct Submission
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24819576.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clones and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GHPF
 Center clone name: CH230-66D7
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 218312 bases at least Q40
 Consensus quality: 223611 bases at least Q30
 Consensus quality: 226962 bases at least Q20
 Estimated insert size: 230141; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved

1 214115: contig of 214115 bp in length

* 214116 214215: gap of unknown length
 * 214116 246644: contig of 32429 bp in length
 * 246644 246744: gap of unknown length
 * 246744 248039: contig of 1295 bp in length
 * 248039 248139: gap of unknown length
 * 248139 249197: contig of 1058 bp in length
 * 249197 249297: gap of unknown length
 * 249297 250782: contig of 1485 bp in length
 * 250782 250882: gap of unknown length
 * 250882 252816: contig of 1934 bp in length
 * 252816 252916: gap of unknown length
 * 252916 254581: contig of 1665 bp in length.
 Location/Qualifiers
 1..254581
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-66D7"
 1343..1992
 /note="clone boundary
 clone end:T7
 site:EcoRI
 end sequence:BH264958"
 3545..5257
 /note="wgs contig"
 134882..138750
 /note="wgs contig"
 162643..165387
 /note="wgs contig"
 195388..196928
 /note="wgs contig"

ORIGIN
 Query Match 1.8%; Score 21; DB 2; Length 254581;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 477 CCATGGAAGCAGAGAGAGCA 497
 |||||
 Db 39422 CCATGGAAGCAGAGAGAGCA 39402
 |||||

RESULT 74
 AC129130 251476 bp DNA linear HTG 21-SEP-2002
 LOCUS Rattus norvegicus clone CH230-107E22, *** SEQUENCING IN PROGRESS
 DEFINITION *** 2 unordered pieces.
 AC129130
 Rattus norvegicus
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)
 SOURCE
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 251476)
 Muzny, D., Marie, J., Mettler, M., Lee, J., Adams, C., Alder, J.,
 Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaru, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bialvo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,
 Egan, A., Escotto, M., Evans, C., Evans, C., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorjis, E., Geer, K., Gill, R., Grady, M., Guerra, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,


```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 257380: contig of 257380 bp in length
* 257381 257480: gap of unknown length
* 257481 261476: contig of 3996 bp in length.

FEATURES
    source
        1..261476
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-107E22"
            1..1222
            /notes="wgs end extension"
            clone end: T7
            complement(7969..8880)
            /note="clone boundary"
            clone end: T7
            site: EcoRI
            end sequence: BH359430"
            193323..195461
            /note="wgs_contig"

    misc_feature
        1..1222
            /notes="wgs end extension"
            clone end: T7
            complement(7969..8880)
            /note="clone boundary"
            clone end: T7
            site: EcoRI
            end sequence: BH359430"
            193323..195461
            /note="wgs_contig"

    misc_feature
        193323..195461
            /note="wgs_contig"

ORIGIN
Query Match      1.8%; Score 21; DB 2; Length 261476;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 586 TGAGGATGGAGACTGGTGGAC 606
Db 95219 TGAGGATGGAGACTGGTGGAC 95239

AC115346 270006 bp DNA linear HTG 13-NOV-2002
Rattus norvegicus clone CH230-96115, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC115346 GI:24941796
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 270006)
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, T., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
Karpaty, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kwis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshevari, L., Loulaeged, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawiney, S., McLeod, M. P., McNell, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G. G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokelemeh, O., Okwuonu, G., Olarpunsaagoon, A., Pal, S., Parke, K.,
Pasternak, S., Paul, H., Perez, A., Perez, E., Pfankech, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Fu, L. L.,
Pudzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajd, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waidron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willison, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, J., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 261476)
Worley, K. C.
Direct Submission
Submitted (27-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 261476)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 21, 2002 this sequence version replaced gi:21998916.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSWC
Center clone name: CH230-107E22
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 239408 bases at least Q40
Consensus quality: 241615 bases at least Q30
Consensus quality: 243058 bases at least Q20
Estimated insert size: 260468; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.

```

COMMENT

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTKN
Center clone name: CH230-96115
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 258636 bases at least Q40
Consensus quality: 25915 bases at least Q30
Consensus quality: 261246 bases at least Q20
Estimated insert size: 264905; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.bcm.tmc.edu/docs/genbank\_draft\_data.html)

```

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angulano, D., Anylebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benamed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, C., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A. D. Souza, B.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowls, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisedge, H., Lozado, R.J., Lu, K., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montanavor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Pasarek, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shivartsbeyn, A., Sibson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, A., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 272636)
 Worley, K.C.
 Direct Submission
 Submitted (31-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 272636)
 Rat Genome Sequencing Consortium.
 Submitted (22-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Sep 22, 2002 this sequence version replaced gi:21737791.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center of Medicine
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GPSZ
 Center clone name: CH230-199K6
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 242962 bases at least Q40

Consensus quality: 246182 bases at least Q30
 Consensus quality: 248330 bases at least Q20
 Estimated insert size: 262434; sum-of-contigs estimation
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 10583: contig of 10583 bp in length
 * 10584 10683: gap of unknown length
 * 10684 267966: contig of 257283 bp in length
 * 267967 268066: gap of unknown length
 * 268067 271169: contig of 3103 bp in length
 * 271170 271269: gap of unknown length
 * 271270 272636: contig of 1367 bp in length.

FEATURES
 source
 1. 272636
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /chromosome="18"
 /clone="CH230-199K6"
 1. 1338
 /notes="wgs end extension
 clone end:Sp6"
 4198..4624
 /notes="clone boundary
 clone end:Sp6
 site:ECORI
 end sequence:BH265145"
 10684..13454
 /notes="wgs contig"
 22942..24018
 /notes="wgs_contig"

misc_feature
 misc_feature
 misc_feature
 misc_feature

Query Match 1.8%; Score 21; DB 2; Length 272636;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1135 CTCTCATCAGCCTGAATGA 1155
 |||||
 Db 210661 CTCTCATCAGCCTGAATGA 210681

RESULT 77
 AC105843/c
 LOCUS
 DEFINITION
 Rattus norvegicus chromosome 18 clone CH230-33G4, *** SEQUENCING IN PROGRESS ***
 AC105843
 ACCESSION
 VERSION
 AC105843.8 GI:23265684
 HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS
 Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM
 Rattus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 337732)
 Muzny, D., Marie, Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, P., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregergis, E., Geer, K., Gill, R., Grady, M., Guerra, T., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenschewa, L., Louisedge, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwankwemeh, O., Okwuonu, G., Olampunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, J., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 337732)
Worley, K.C.

Direct Submission
Submitted (10-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 337732)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 21, 2002 this sequence version replaced gi:21736972.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMD
Center clone name: CH230-33G4

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 234247 bases at least Q40
Consensus quality: 235102 bases at least Q30
Consensus quality: 235456 bases at least Q20
Estimated insert size: 260117; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 337732: Contig of 337732 bp in length.

FEATURES
source
1. 337732
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/chromosome="18"
/clone="CH230-33G4"
1. 1114
/notes="wgs_end_extension
clone_end:Sp6"
74166..171591
/notes="clone_boundary
clone_end:Sp6"
site:EcoRI
end_sequence:BH353835"
complement(237325..238035)
/notes="clone_boundary
clone_end:T7"
site:EcoRI
end_sequence:BH353833"
336434..337732
/notes="wgs_end_extension
clone_end:T7"

misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Query Match 1.8%; Score 21; DB 2; Length 337732;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1135 CTTCTACATCAGCTGAATGA 1155
Db 176040 CTTCTACATCAGCTGAATGA 176020

RESULT 78
AX443199/c
LOCUS AX443199 20 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 140 from Patent WO0216599.
ACCESSION AX443199
VERSION AX443199.1 GI:21690594
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R., Shinkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E., Topper, J.N. and Yang, R.B.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0216599-A 140 28-FEB-2002;
Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES Location/Qualifiers

```

source      1..20
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="oligonucleotide primer"

ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 282 CAAGCCCTAACCTGCTCCAG 301
Db 20 CAAGCCCTAACCTGCTCCAG 1

RESULT 79
LOCUS AX452883/c 20 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 4 from Patent WO0242457.
ACCESSION AX452883
KEYWORDS AX452883.1 GI:21712522
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1
AUTHORS Chang,H., Yang,W.P., Wu,Y., Whitney,G.S., Perez-Villar,J.J. and
          Kanner,S.B.
TITLE Cloning and expression of human slap-2: a novel sh2/sh3
          domain-containing human slap homologue having immune cell-specific
          expression
JOURNAL Patent: WO 0242457-A 4 30-MAY-2002;
          Bristol-Myers Squibb Co. (US)
FEATURES
source      1..20
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="FPY751 PCR PRIMER"

ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1138 CTACATCAGCCTGAATGACG 1157
Db 20 CTACATCAGCCTGAATGACG 1

RESULT 80
LOCUS AX909683/c 277 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 25546 from Patent EP1033401.
ACCESSION AX909683
VERSION AX909683.1 GI:40065763
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 25546 06-SEP-2000;
          Genset (FR)
FEATURES
source      1..277
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 277;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 AGATCCTCCAGGCTGAGAG 229
Db 221 AGATCCTCCAGGCTGAGAG 202

RESULT 81
LOCUS BD045216/c 277 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD045216
VERSION BD045216.1 GI:22586958
KEYWORDS JP 2001269182-A/21462.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 277)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 21462 02-OCT-2001;
          GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/21462
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PI 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
PC C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
CC G06F15/40

FEATURES
source      1..277
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 277;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 AGATCCTCCAGGCTGAGAG 229
Db 221 AGATCCTCCAGGCTGAGAG 202

RESULT 82
LOCUS AX867799 525 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 2704 from Patent EP1074617.
ACCESSION AX867799
VERSION AX867799.1 GI:40022662
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
          Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 2704 07-FEB-2001;
          Research Association for Biotechnology (JP)
FEATURES
source      Location/Qualifiers

```

```

source
1. .525
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 1.7%; Score 20; DB 6; Length 525;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 GAAGAAATCTCTGCCAAGC 436
DB 452 GAAGAAATCTCTGCCAAGC 471

RESULT 83
BD147861 525 bp DNA linear PAT 17-JAN-2003
LOCUS
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
ACCESSION
BD147861
VERSION
BD147861.1 GI:27853619
KEYWORDS
JP 2002191363-A/2704.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 525)
AUTHORS
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE
Primer for synthesizing full-length cDNA and use thereof.
JOURNAL
Patent: JP 2002191363-A 2704 09-JUL-2002;
HELEX RESEARCH INSTITUTE
COMMENT
OS Homo sapiens (human)
PN JP 2002191363-A/2704
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/10, C12P21/02, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT source 1. .525
FT Location/Qualifiers
/organism="Homo sapiens (human)".
FEATURES
source
1. .525
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 1.7%; Score 20; DB 6; Length 525;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 GAAGAAATCTCTGCCAAGC 436
DB 452 GAAGAAATCTCTGCCAAGC 471

RESULT 84
AR418884 539 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION
Sequence 10381 from patent US 6639063.
ACCESSION
AR418884
VERSION
AR418884.1 GI:40173994
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unknown.

Unclassified.
1 (bases 1 to 539)
Edwards,J.B.D.M., Jobert,S. and Giordano,J.-Y.
AUTHORS
EST's and encoded human proteins
TITLE
Patent: US 6639063-A 10381 28-OCT-2003;
JOURNAL
Location/Qualifiers
FEATURES
source
1. .539
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 1.7%; Score 20; DB 6; Length 539;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 TGTGACCATGGAAGCAGAGA 491
DB 52 TGTGACCATGGAAGCAGAGA 71

RESULT 85
BD114437 539 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION
EST and encoded human protein.
ACCESSION
BD114437
VERSION
BD114437.1 GI:23209341
KEYWORDS
JP 2002010789-A/6514.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 539)
AUTHORS
Edwards,J.B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE
EST and encoded human protein
JOURNAL
Patent: JP 2002010789-A 6514 15-JAN-2002;
GENSET CORP
COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/6514
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PI 05-AUG-1999 US 60/47499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE
PI GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT source 1. .539
FT Location/Qualifiers
/organism="Homo sapiens (human)".
FEATURES
source
1. .539
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 1.7%; Score 20; DB 6; Length 539;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 TGTGACCATGGAAGCAGAGA 491
DB 52 TGTGACCATGGAAGCAGAGA 71

RESULT 86
AR016398 675 bp DNA linear PAT 05-DEC-1998
LOCUS
DEFINITION
Sequence 3 from patent US 5776696.
ACCESSION
AR016398
KEYWORDS
Accession
SOURCE
Accession
VERSION
AR016398.1 GI:3972675

```

```

KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 675)
AUTHORS     Salowe,S.P.
TITLE       High throughput assay using fusion proteins
JOURNAL     Patent: US 5776696-A 3 07-JUL-1998;
FEATURES    Location/Qualifiers
             source
             1..675
             /organism="unknown"
             /mol_type="unassigned DNA"
ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 675;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCTCATCGGAGAGC 766
Db 443 CCTTCTCATCGGAGAGC 462

RESULT 87
LOCUS      AR019256          675 bp    DNA    linear    PAT 05-DEC-1998
DEFINITION Sequence 3 from patent US 5783398.
ACCESSION  AR019256
VERSION     AR019256.1 GI:3974370
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 675)
AUTHORS     Marcy,A., Salowe,S.P. and Wisniewski,D.
TITLE       High throughput assay using fusion proteins
JOURNAL     Patent: US 5783398-A 3 21-JUL-1998;
FEATURES    Location/Qualifiers
             source
             1..675
             /organism="unknown"
             /mol_type="unassigned DNA"
ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 675;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCTCATCGGAGAGC 766
Db 443 CCTTCTCATCGGAGAGC 462

RESULT 88
LOCUS      AX653434          1047 bp    DNA    linear    PAT 22-MAR-2003
DEFINITION Sequence 3304 from Patent WO03000898.
ACCESSION  AX653434
VERSION     AX653434.1 GI:29156248
KEYWORDS
SOURCE      Oryza sativa
ORGANISM    Oryza sativa
REFERENCE   1
AUTHORS     Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M., Katagiri,P., Qian,Y., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE       Plant genes involved in defense against pathogens
JOURNAL     Patent: WO 03000898-A 3304 03-JAN-2003;
FEATURES    Syngenta Participations AG (CH)
             Location/Qualifiers
             source
             1..1047
             /organism="Oryza sativa"

KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 675)
AUTHORS     Salowe,S.P. and Wisniewski,D.
TITLE       High throughput assay using fusion proteins
JOURNAL     Patent: US 5783398-A 3 21-JUL-1998;
FEATURES    Location/Qualifiers
             source
             1..675
             /organism="unknown"
             /mol_type="unassigned DNA"
ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 675;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCTCATCGGAGAGC 766
Db 443 CCTTCTCATCGGAGAGC 462

RESULT 89
LOCUS      AF228313          1491 bp    mRNA    linear    PRI 17-FEB-2000
DEFINITION Homo sapiens tyrosine kinase LCK mRNA, partial cds.
ACCESSION  AF228313
VERSION     AF228313.1 GI:6984208
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1491)
AUTHORS     Boncrisiano,M., Majolini,M.B., D'Ellos,M.M., Pacini,S., Valensin,S., Ulivieri,C., Falini,B., Del Prete,G., Telford,J.L. and Baldari,C.T.
TITLE       Defective recruitment and activation of ZAP-70 in CVID patients with T-cell defects
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1491)
AUTHORS     Boncrisiano,M. and Baldari,C.T.
TITLE       Direct Submission
JOURNAL     Submitted (25-JAN-2000) Department of Evolutionary Biology, University of Siena, Via Mattioli 4, Siena 53100, Italy
FEATURES    Location/Qualifiers
             source
             1..1491
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /tissue_type="peripheral blood lymphocytes"
             <1..1491
             /codon_start=1
             /product="tyrosine kinase LCK"
             /protein_id="AAF34794.1"
             /db_xref="GI:6984209"
             /translation="MENIDVCNCHYPVPLDQKGTLLIRNGSEVRDPLVTEGNSNPP
             ASPLQNLVIALHSYEPHSDGDLGPEKQLRLEQSGEWKQAQSLTTQGEPIPNF
             VAKANSLPEPWFKNLSKDAERQLAFNGNTHSGFLRESSTAGSFLSVRFDFQDN
             QGEVVRHYKIRLNDNGFYISPRITPPGLHVLRYHTNASDGLCTLSRQCQKQPK
             PWEDEWEVPRETLKLVERLGAGQFGEVWVGNGHTKVAVKSLKQGSMPDAFLAEA
             NLMKQLHORLVRLVAVTQEPYIITEYMENGLVDPLKTPSGIKLTINKLIDMAAO
             IAEGMAFIERNYIHRDLRAAILVSDTLSCIADPGLARLIEDNEYTAREGAKFPK
             WTAPEALNTGTFIKSDVWSFGILLTEIVHGRIPYIPYPMTEVIONLERGVNWRPD
             NCPEEDYQLMRLCWKPERPEDPTFDYLRSLVEDFFFTATEGYQYQP"
ORIGIN
Query Match      1.7%; Score 20; DB 9; Length 1491;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCTCATCGGAGAGC 766
Db 410 CCTTCTCATCGGAGAGC 429

RESULT 90
LOCUS      SSC277921          1527 bp    mRNA    linear    PRI 20-SEP-2001
DEFINITION Saimiri sciureus partial mRNA for Lck tyrosine kinase (lck gene).
ACCESSION  AJ277921
VERSION     AJ277921.1 GI:14140182
FEATURES
SOURCE

```


lck gene; LCK tyrosine kinase.
 Saimiri sciureus (common squirrel monkey)
 Saimiri sciureus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae;
 Saimiri.
 1
 Greve, T., Tanguay, G., Fleischer, B., Fickenscher, H. and Broker, B.M.
 Downregulation of p56(lck) tyrosine kinase activity in T cells of
 squirrel monkeys (Saimiri sciureus) correlates with the
 nontransforming and apathogenic properties of herpesvirus saimiri
 in its natural host
 J. Virol. 75 (19), 9252-9261 (2001)
 MEDLINE 21424508
 PUBMED 11533187
 REFERENCE 2 (bases 1 to 1527)
 Fickenscher, H.
 Direct Submission
 Submitted (15-MAY-2000) Fickenscher H., Institut fuer Klinische und
 Molekulare Virologie, Friedrich-Alexander-Universitaet
 Erlangen-Nuernberg, Schlossgarten 4, D-91054 Erlangen, GERMANY
 Location/Qualifiers
 FEATURES
 source
 1..1527
 /organism="Saimiri sciureus"
 /mol_type="mRNA"
 /db_xref="taxon:9521"
 /cell_type="T lymphocytes"
 1..1527
 /gene="lck"
 1..>1527
 /gene="lck"
 /note="Src-family tyrosine kinase Lck"
 /codon_start=1
 /evidence=experimental
 /product="Lck tyrosine kinase"
 /protein_id="CAC38871.1"
 /db_xref="GI:14140183"
 /db_xref="GI:14140183"
 /db_xref="GGA:Q95KR7"
 /db_xref="SPTREMBL:Q95KR7"
 /translation="MGCGSSHPEDDMENIDVCNCHYPIVPLDCKATLLFRNGSEV
 RDLVRYEGSNPPASPLQDNLVIALHSYEPDGDGLGPEKGEHLRILEQNGWKAQS
 LTTQGEVFPFNFAKANSLEPEPFNKLGRDAERQLLAPGNTHGSLFIRESESTA
 GSFSLVDFDQNGEVVVKYIRNLNDGGFYISPRITFSGHLVRYHTNASDGLCT
 RLSRCPQTKQPKWEDEWEPRTKLVERLGAGQFGEVMGYNEHTKVAKSLK
 QGSMSPDFALENLKOLQKLVRLVAVVTESPIVITEVMENGLVDFLTPSGI
 KLTINKLMDAQIVEGMAFEERNYIHRDLRAANILVSDTSLCKIADFLGLRIEDN
 EYTBEGAKFPFKWTAPEAINYGTPTIKSDVMSFGILMTIELVTGRIPFGMTNPEVI
 QNLRGYRMPRPNCPPELYQLMMLCWKEPDRPTFDYLRSLVLEDDFFFTATEGQYQFQ
 P"
 ORIGIN
 Query Match 1.7%; Score 20; DB 9; Length 1527;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 747 CCTTCTCATCCGGAGAGC 766
 Db 449 CCTTCTCATCCGGAGAGC 468
 RESULT 91
 AX695859
 LOCUS 1530 bp DNA linear PAT 31-MAR-2003
 DEFINITION Sequence 1486 from Patent WO03008583.
 ACCESSION AX695859
 VERSION AX695859.1 GI:29419024
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Morris, D.W. and Engelhard, E.K.

TITLE Novel compositions and methods for cancer
 JOURNAL Patent: WO 03008583-A 1486 30-JAN-2003;
 Sagres Discovery (US)
 FEATURES
 source
 1..1530
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 1.7%; Score 20; DB 6; Length 1530;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 747 CCTTCTCATCCGGAGAGC 766
 Db 449 CCTTCTCATCCGGAGAGC 468
 RESULT 92
 HSP320182
 LOCUS 1530 bp mRNA linear PRI 19-JUN-2003
 DEFINITION Hylobates sp. mRNA for lck protein.
 ACCESSION AJ320182
 VERSION AJ320182.1 GI:14627117
 KEYWORDS lck gene; lck protein; tyrosine kinase.
 SOURCE Hylobates sp. (Gibbon)
 ORGANISM Hylobates sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 1
 Picard, C.
 AUTHORS
 JOURNAL Department of Experimental Oncology laboratory, U
 REFERENCE 2
 Picard, C., Greenway, A., Holloway, G., Olive, D. and Collette, Y.
 Interaction with simian Hck tyrosine kinase reveals convergent
 evolution of the Nef protein from simian and human immunodeficiency
 viruses despite differential molecular surface usage
 Virology 295 (2), 320-327 (2002)
 JOURNAL 22031236
 MEDLINE 12033791
 PUBMED
 REFERENCE 3 (bases 1 to 1530)
 Picard, C.
 AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (02-JUL-2001) Picard C., U119, Inserm, bd Lei Roure,
 Marseille 13010, FRANCE
 FEATURES
 source
 1..1530
 /organism="Hylobates sp."
 /mol_type="mRNA"
 /db_xref="taxon:9581"
 1..1530
 /gene="lck"
 1..1530
 /gene="lck"
 /function="tyrosine kinase"
 /codon_start=1
 /product="lck protein"
 /protein_id="CAC44027.1"
 /db_xref="GI:14627118"
 /db_xref="GGA:Q95M32"
 /db_xref="SPTREMBL:Q95M32"
 /translation="MGCGSSHPEDDMENIDVCNCHYPIVPLDCKATLLFRNGSEV
 RPLVRYEGSNPPASPLQDNLVIALHSYEPDGDGLGPEKGEHLRILEQNGWKAQS
 LTTQGEVFPFNFAKANSLEPEPFNKLGRDAERQLLAPGNTHGSLFIRESESTA
 GSFSLVDFDQNGEVVVKYIRNLNDGGFYISPRITFSGHLVRYHTNASDGLCT
 RLSRCPQTKQPKWEDEWEPRTKLVERLGAGQFGEVMGYNEHTKVAKSLK
 QGSMSPDFALENLKOLQKLVRLVAVVTESPIVITEVMENGLVDFLTPSGI
 KLTINKLMDAQIVEGMAFEERNYIHRDLRAANILVSDTSLCKIADFLGLRIEDN
 EYTBEGAKFPFKWTAPEAINYGTPTIKSDVMSFGILMTIELVTGRIPFGMTNPEVI
 QNLRGYRMPRPNCPPELYQLMMLCWKEPDRPTFDYLRSLVLEDDFFFTATEGQYQFQ
 P"
 ORIGIN


```

QY 747 CCTTCTCATCCGGGAGAC 766
    |||||||
Db 475 CCTTCTCATCCGGGAGAC 494

RESULT 95
AY335586 1620 bp mRNA linear SYN 15-OCT-2003
LOCUS Synthetic construct Homo sapiens lymphocyte-specific protein
DEFINITION tyrosine kinase (LCK) mRNA, partial cds.
ACCESSION AY335586
VERSION AY335586.1 GI:33303798
KEYWORDS FLI CDNA.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1620)
AUTHORS Park,J., Rolf,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D.,
Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepson,D., Harlow,E.,
LaBaer,J. and Brizuela,L.
TITLE Cloning of human full-length CDS FLEXGene kinases in
recombinational vector system
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1620)
Park,J., Rolf,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D.,
Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepson,D., Harlow,E.,
LaBaer,J. and Brizuela,L.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2003) Biological Chemistry and Molecular
Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
Cambridge, MA 02141-2023, USA
COMMENT This CDS clone is a part of a collection of human full-length
expression clones generated by Harvard Institute of Proteomics.
Each CDS has been cloned without stop-codon (to allow fusion with
C-terminal tag). The CDS has been directionally cloned using BD
In-Fusion(TM) cloning system between the Sali and HindIII sites of
the pDNR-Dual vector. Additional sequences in the clone: 'ACC'
after Sali site and before 'ATG' to provide Kozak consensus
sequence; 'GG' after last codon and before HindIII site to maintain
reading frame.
FEATURES
    Location/Qualifiers
        1..1620
            /organism="synthetic construct"
            /mol_type="mRNA"
            /db_xref="taxon:32630"
            /clone_lib="MGC clone templates"
            /lab_host="DH5alpha T1 resistant"
            /note="Vector: pDNR-Dual"
        1..>1620
            /gene="LCK"
        1..>1620
            /gene="LCK"
            /note="Mutations: Stop->Leu"
            /codon_start=1
            /transl_table=11
            /product="lymphocyte-specific protein tyrosine kinase"
            /protein_id="AA02413.1"
            /db_xref="GI:33303798"
            /translation="MGCSSHPDDWNIIDVCNCHYPIVPIVDGKGTLLINGSEV
RDPLVTGSPNPASPQDNILVIALHSYPSHDGLDFEKGQRLIEQSGEWMKQAS
LTTQEGIPFNFVAKNSLPFEPFNLSRKDAERQLLAPGNTHGSLFIRSESTA
GSFSLVSVDQFQNGEVVVKYKIRNLNGGFIYISPRITFGHLVHGYTHNSDGLCT
RLSRPCOTKPKDWEVETLKLVERLGAGOFGEVVMGYNGHKKVAKSLK
QGSMPDAFLAENLQLOHRLVLYAVVTOEPIYLIYEYMDNDFLLSQLEEKL
CASPWNLGQQQLLLPTGSLVDVFKTPSGIKLTKLLDMAAQIAEGMAIERNYIH
RDLRAANLIVSDTUSKIADEGLARLIEDNEYTAREGKFKPMTAPEALNYGFTTIK
SDVMSFGLLLTIVTHGRIPYPMNTNPENLGRGYRVRPDCNPCELYQLARLCKW
ERPEDRPTFDYLRSLVDFFFTATEGQYQPOPL"

Query Match 1.7%; Score 20; DB 12; Length 1620;

```

```

Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGAC 766
    |||||||
Db 449 CCTTCTCATCCGGGAGAC 468

RESULT 96
AK114952/c
LOCUS AK114952 1707 bp mRNA linear INV 30-NOV-2002
DEFINITION Ciona intestinalis cDNA, clone:cieg024d02, full insert sequence.
ACCESSION AK114952
VERSION AK114952.1 GI:23586708
KEYWORDS FLI CDNA.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
REFERENCE 1
AUTHORS Satou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawashima,T.,
Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasakura,Y.,
Nakayama,A., Ishikawa,H., Inaba,K. and Satoh,N.
TITLE A cDNA resource from the basal chordate Ciona intestinalis
JOURNAL Genesis 33 (4), 153-154 (2002)
MEDLINE 22191024
FURMED 12203911
REFERENCE 2 (bases 1 to 1707)
Satou,Y. and Satoh,N.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Fax:81-75-705-1113
COMMENT Ciona intestinalis cDNA Project (URL:
http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
FEATURES
    Location/Qualifiers
        1..1707
            /organism="Ciona intestinalis"
            /mol_type="mRNA"
            /db_xref="taxon:7719"
            /clone="cieg024d02"

ORIGIN

Query Match 1.7%; Score 20; DB 3; Length 1707;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 847 CCACCTGCCTTGACATGGCT 866
    |||||||
Db 936 CCACCTGCCTTGACATGGCT 917

RESULT 97
AX695858
LOCUS AX695858 2032 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 1485 from Patent WO03008583.
ACCESSION AX695858
VERSION AX695858.1 GI:29419023
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Morris,D.W. and Engelhard,E.K.
TITLE Novel compositions and methods for cancer
JOURNAL Patent: WO 03008583-A 1485 30-JAN-2003;
Sagres Discovery (US)
FEATURES
    Location/Qualifiers
        1..2032
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"

ORIGIN

```

```

/db_xref="taxon:9606"

ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 2032;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCCTCATCCGGAGAGC 766
Db 500 CCTTCCTCATCCGGAGAGC 519

RESULT 99
BD223388
LOCUS      2032 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Method of changing protein activity in PKA signaling pathway.
ACCESSION  BD223388
VERSION     BD223388.1 GI:33033158
KEYWORDS    JP 2002516670-A/3.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 2032)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       1 (bases 1 to 2032)
JOURNAL     Hansson V., Levy F.O., Mustelin T., Skalhegg B.S., Sundvold V.,
            Tasken K., Vang T., Altman A. and Munsni A.
METHOD      Method of changing protein activity in PKA signaling pathway
PATENT      Patent: JP 2002516670-A 3 11-JUN-2002;
            LAURAS AS
COMMENT     OS Homo sapiens (human)
            PN JP 2002516670-A/3
            PD 11-JUN-2002
            PF 27-MAY-1999 JP 2000551593
            PR 27-MAY-1998 NO 19982419,30-DEC-1998 US 60/114240 PI
            VIDAR HANSSON, FINN OLAV LEVY, THOMAS MUSTELIN, BJORN STEEN PI
            SKALHEGG,
            PI VIBEKE SUNDVOLD, KJETIL TASKEN, TORIKEL VANG, AMNON ALTMAN, ANIL
            MUNSHI
            PC C12N15/09,A61K38/00,A61K45/00,A61K48/00,A61P35/00,A61P37/02,
            A61P43/00,
            PC C07K14/47,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12Q1/02 PC
            C12N15/00,A61K37/02,
            PC C12N5/00
            CC Method of changing protein activity in PKA signaling pathway
            FH Key
            FT CDS
            Location/Qualifiers
            1..2032
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      1.7%; Score 20; DB 6; Length 2032;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCCTCATCCGGAGAGC 766
Db 500 CCTTCCTCATCCGGAGAGC 519

RESULT 99
HSLCKB
LOCUS      2032 bp      mRNA      linear      PRI 12-SEP-1993
DEFINITION Human lck mRNA for membrane associated protein tyrosine kinase.
ACCESSION  X13529
VERSION     X13529.1 GI:34294
KEYWORDS    lck gene; membrane-associated protein; oncogene; phosphoprotein;
            tyrosine protein kinase.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 2032)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       1 (bases 1 to 2032)
JOURNAL     Perlmutter R.M., Marth J.D., Lewis D.B., Peet R., Ziegler S.F. and
            Wilson C.B.
METHOD      Structure and expression of lck transcripts in human lymphoid cells
PATENT      J. Cell. Biochem. 38 (2), 117-126 (1988)
COMMENT     source text: Human adult peripheral blood lymphocyte, cDNA
            clones HK[28,8,2,5].
            to mRNA,
            Location/Qualifiers
            1..2032
            /organism="Homo sapiens"
            /mol_type="mRNA"

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2032)
Perlmutter R.M., Marth J.D., Lewis D.B., Peet R., Ziegler S.F. and
Wilson C.B.
Structure and expression of lck transcripts in human lymphoid cells
J. Cell. Biochem. 38 (2), 117-126 (1988)
MEDLINE 89123626
PUBMED 3265417
COMMENT See <X06369> for overlapping sequence.
FEATURES
Location/Qualifiers
1..2032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="lp32-35"
/cell_type="leukocyte"
52..1581
/note="unnamed protein product; lck protein (AA 1-509)"
/codon_start=1
/protein_id="CAA31884.1"
/db_xref="GI:34295"
/db_xref="GOA:P06239"
/db_xref="SWISS-PROT:P06239"
/translation="MGCGSSHPEDDWMENIDVCNCHYPIVPLDCKGTLIRNGSEV
RPLVTYEGSNPPASPLQDNLVIALHSYSPSHDGLGFKGEPKPLRILEQSGWKAQS
LITGQGFIPFNFAKANSLEPEKFNLSKDAEROLAPGNTHGSLRSESTA
GAFSLGVRFDQNGQGVYKIKNLNGGFEYISPRITPGGHELVRYHTWASDGLCT
RUSRFCTQKPKQWDEWEVRETLLKVLKRGAGQGVGVNGYNGHTKVAVSLK
QGSMSDFALAEANLKLQHQRLVLYAVTQEPYIITEYMENGLVDFDKTSGI
KLITNKLDMAAQIAEGNAFIEERNYIHRDLAAANILVSDTSLCKIADFGLARLIEDN
EYTAGAKFPKWTAPALNYGTFTIKSDVMSFGILLTEIVTHGRIPYPGMTNPEVI
ONLERYVMVDPDNCPEELYQLMRLCWKPERDPTFDYLRVSLVLEDFFTATEGQVQ
P"
misc_feature 2008..2013
            /note="polyadenylation signal"
polyA_site 2032
            /note="polyadenylation site"

ORIGIN
Query Match      1.7%; Score 20; DB 9; Length 2032;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCCTCATCCGGAGAGC 766
Db 500 CCTTCCTCATCCGGAGAGC 519

RESULT 100
HUMLCCKAA
LOCUS      2032 bp      mRNA      linear      PRI 07-JAN-1995
DEFINITION Human lymphocyte-specific protein tyrosine kinase (lck) mRNA,
            complete cds.
ACCESSION  M36881
VERSION     M36881.1 GI:187033
KEYWORDS    lymphocyte-specific protein tyrosine kinase.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 2032)
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Perlmutter R.M., Marth J.D., Lewis D.B., Peet R., Ziegler S.F. and
            Wilson C.B.
TITLE       1 (bases 1 to 2032)
JOURNAL     Structure and expression of lck transcripts in human lymphoid cells
            J. Cell. Biochem. 38 (2), 117-126 (1988)
MEDLINE 89123626
PUBMED 3265417
COMMENT Original source text: Human adult peripheral blood lymphocyte, cDNA
            clones HK[28,8,2,5].
            to mRNA,
            Location/Qualifiers
            1..2032
            /organism="Homo sapiens"
            /mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/map="lp35-p32"
1..2032
/gene="LCK"
mRNA
<1..2032
/gene="LCK"
/feature="lymphocyte-specific tyrosine kinase; G00-119-360"
52..1581
/gene="LCK"
/codon_start=1
/product="lymphocyte-specific protein tyrosine kinase"
/protein_id="AAA59502.1"
/db_xref="GI:187034"
/db_xref="GDB:G00-119-360"
/translation="MGCCSSHPEDDMENIDVCENCHYPIVRLDGKRLIRNGSEV
RDPLVYEGSNPPASPLQDNLVIALHSYPSHDGLGFERKEPLRILEQSGEWWKAQS
LTTGQEGFIPFNFAKASLEPEPFKNLSRKAERQLAPGNTHGSLIRSESTA
GSFSLVDRDFQNGEVVYKIRNLONGFYISPRITFGHLEHVRHYTNASDGLCT
RLSRPCOTKPKQKWEDEWVPRTELKVERLGAOGEVWVGYYNGHTKVAVKSILK
QGSMSPDALAEANLMKQLQHQRLVRLYAVVTQBPYIITTEYMENGSLVDFLTPSGI
KLITNKLDMAQAAGMAFTEERNYIHRDURANILVSDLSCKIADFGRLIEDN
EYTAGAKFPIKWTAPAINYGTFTIKSDVMSFGILLTEIVTHGRIPYGMNTPEVI
QNLEGYRMVRPDCPELYQLMLRCWKERPEDRPTFDYLRSLVLEDFTTATEGQYQFQ
P"

```

ORIGIN

```

Query Match      1.7%; Score 20; DB 9; Length 2032;
Best Local Similarity 100.0%; Pred.No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      747 CCTTCCTCATCGGAGAGC 766
          |||||
Db       500 CCTTCCTCATCGGAGAGC 519

```

Search completed: March 25, 2004, 07:38:41
Job time : 4892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 03:56:37 ; Search time 531 Seconds
(without alignments)
9464.441 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183
Sequence: 1 agctagagctccaggaacc.....tctcttgatgatgcctag 1183

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database : N Geneseq 29Jan04:*

- 1: Geneseqm1980s:*
- 2: Geneseqm1990s:*
- 3: Geneseqm2000s:*
- 4: Geneseqm2001as:*
- 5: Geneseqm2001bs:*
- 6: Geneseqm2002as:*
- 7: Geneseqm2003as:*
- 8: Geneseqm2003bs:*
- 9: Geneseqm2003cs:*
- 10: Geneseqm2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1183	100.0	1183	6	Abk61465 Human cDN
2	1132	95.7	2567	6	Ad43980 Human Src
3	735	62.1	786	6	Ala44089 Human mod
4	735	62.1	786	6	Abq74343 Human Src
5	657	55.5	763	6	Abq98670 Human ORF
6	592	50.0	837	3	Aac77202 Human ORF
7	483	40.8	737	6	Ala44090 Mouse MAR
8	468	39.6	1413	6	Abq99374 Human cod
9	405	34.2	603	5	Aas74748 Human cod
10	401	33.9	2049	5	Aas74750 DNA encod
11	348	29.4	444	6	Abq98669 Human ORF
12	341	28.8	875	6	Abq99151 Human ORF
13	244	20.6	445	5	Aas74747 Human ORF
14	141	11.9	211	5	Aas70181 DNA encod
15	134	11.3	432	4	Ala12879 Probe #28
16	134	11.3	432	4	Aas4580 Human foe
17	134	11.3	432	4	Ala34236 Probe #29
18	134	11.3	432	4	Ala44128 Human bre
19	134	11.3	432	4	Ala24363 Probe #28
20	134	11.3	432	4	Aak28314 Human bon
21	134	11.3	432	4	Aak02872 Human bra
22	134	11.3	432	4	Abq27912 Human liv
23	134	11.3	432	5	Ala102797 Probe #27

24 134 11.3 432 6 ABS02823 Human gen
25 134 11.3 448 4 Aal14520 Probe #44
26 134 11.3 448 6 ABS04499 Human gen
27 96 8.1 96 4 Aal22119 Probe #12
28 96 8.1 96 4 ABA67198 Human foe
29 96 8.1 96 4 AAI47414 Probe #16
30 96 8.1 96 4 ABA49284 Human bre
31 96 8.1 96 4 ABA34292 Probe #12
32 96 8.1 96 4 AAK41374 Human bon
33 96 8.1 96 4 AAK15640 Human liv
34 96 8.1 96 4 ABS40966 Human liv
35 96 8.1 96 5 AAI07818 Probe #78
36 96 8.1 96 6 ABS15380 Human gen
37 43 3.6 1348 6 AAL44087 Mouse mod
38 30 2.5 30 6 AD43983 Human leu
39 26 2.2 26 6 ABK61506 Human NOV
40 25 2.1 25 6 AAL44100 Human mod
41 25 2.1 25 6 AAL44098 Human mod
42 25 2.1 25 6 AAL44099 Human mod
43 23 1.9 23 6 AAL44097 Human mod
44 22 1.9 22 6 ABK61507 Human NOV
45 22 1.9 3070 2 AAX55277 Nucleotid
46 22 1.9 3070 2 AAX55278
47 21 1.8 21 6 AD43981 Human leu
48 20 1.7 20 6 ABK61505 Human NOV
49 20 1.7 20 6 AD43982 Human leu
50 20 1.7 277 3 AAC21471 Human sec
51 20 1.7 525 4 AAH05869 Human cDN
52 20 1.7 675 2 AAT63421 FKBP-LCK:
53 20 1.7 675 2 AAX5151 DNA encod
54 20 1.7 1047 7 ADA69981 Rice gene
55 20 1.7 1254 2 AAI3983 Lck gene
56 20 1.7 1530 8 ADA02968 Human LCK
57 20 1.7 1530 9 AD43982 Human LCK
58 20 1.7 1530 9 ADC85448 Human LCK
59 20 1.7 1928 3 ADC9100 Human KPP
60 20 1.7 2032 3 AAZ46491 PKA subst
61 20 1.7 2032 7 ACC72850 Human can
62 20 1.7 2032 8 ADA02967 Human LCK
63 20 1.7 2032 9 ADB72705 Human LCK
64 20 1.7 2032 9 ADC85447 Human LCK
65 20 1.7 2032 9 ADE40448 Human pro
66 20 1.7 2034 6 ABS65202 cDNA enco
67 20 1.7 2129 7 ACA56854 Human sig
68 20 1.7 2129 7 ACC81082 Human T-1
69 20 1.7 2129 7 ACA64888 Human p56
70 20 1.7 2282 6 ABS65203 cDNA enco
71 20 1.7 2320 5 AAS86451 DNA encod
72 20 1.7 2483 4 AAH46496 Subtilope
73 20 1.7 31842 8 ADA02966 Human LCK
74 20 1.7 31842 9 ADB72704 Human LCK
75 20 1.7 31842 9 ADC85446 Human Lck
76 20 1.7 110000 6 ABQ74964 4 Continuation (5 of
77 19 1.6 65 6 ABN31762 Rat splc
78 19 1.6 114 4 AAK24414 Human bra
79 19 1.6 307 8 ACL21602 DNA clone
80 19 1.6 416 4 AAK83462 Human imm
81 19 1.6 417 4 AAK83463 Human imm
82 19 1.6 417 4 AAK83459 Human imm
83 19 1.6 510 6 ABX04900 Conus sp
84 19 1.6 599 4 AAK11828 Human bra
85 19 1.6 1033 4 AAP91867 Human sec
86 19 1.6 1467 8 ADA09466 Human kin
87 19 1.6 1467 7 ADE38350 Human pro
88 19 1.6 1490 7 ACC57671 Mouse pro
89 19 1.6 1548 2 AAV81743 Human SAD
90 19 1.6 1556 9 ADB53706 Primary r
91 19 1.6 1580 6 ABN86922 Human NOV
92 19 1.6 1761 7 AB234698 Coding se
93 19 1.6 1833 6 ABK89238 Human cDN
94 19 1.6 2120 7 ADA53009 Human KPP
95 19 1.6 2130 10 ADE28356 Human cDN
96 19 1.6 2527 7 ACC46205 Human dit

97	19	1.6	2771	7	ACA56503	Human sig	170	18	1.5	1579	7	ACD18362	Human sec
98	19	1.6	2774	2	AAT60434	Human int	171	18	1.5	1579	7	ACD08369	Human sec
99	19	1.6	2789	2	AAT60433	Human int	172	18	1.5	1579	7	ACA88803	Novel hum
100	19	1.6	3127	5	AAS80650	DNA encod	173	18	1.5	1579	7	ACA70245	Human sec
C 101	19	1.6	3163	4	AAK69686	Human imm	174	18	1.5	1579	7	ACD12467	Novel hum
C 102	19	1.6	3163	4	AAK69685	Human imm	175	18	1.5	1579	7	ACD12467	Novel hum
C 103	19	1.6	110000	7	AD533223_2	Continuation (3 of	176	18	1.5	1579	7	ACD16010	Human sec
C 104	18	1.5	204	5	AAH68226	C glutami	177	18	1.5	1579	7	ACD16010	Human sec
C 105	18	1.5	372	7	ACA01527	H. glutam	178	18	1.5	1579	7	ACD18055	Novel hum
C 106	18	1.5	383	6	ABN23352	Human ORF	179	18	1.5	1579	7	ACD18055	Novel hum
107	18	1.5	402	8	ACH21397	Human adu	180	18	1.5	1579	7	ACD18055	Novel hum
108	18	1.5	403	7	ABX51886	Bovine ES	181	18	1.5	1579	7	ACD18055	Novel hum
109	18	1.5	416	5	AA574360	DNA encod	182	18	1.5	1579	7	ACD18055	Novel hum
110	18	1.5	427	5	ABA12702	Human ner	183	18	1.5	1579	7	ACD18055	Novel hum
111	18	1.5	438	3	AA28674	Human sec	184	18	1.5	1579	7	ACD18055	Novel hum
112	18	1.5	483	7	ABX74718	Human CDN	185	18	1.5	1579	7	ACD18055	Novel hum
113	18	1.5	487	8	ACH39527	Human foe	186	18	1.5	1579	7	ACD18055	Novel hum
114	18	1.5	491	8	ACH50807	Human mam	187	18	1.5	1579	7	ACD18055	Novel hum
115	18	1.5	544	5	ABA06420	Human CDN	188	18	1.5	1579	7	ACD18055	Novel hum
116	18	1.5	544	5	ABA13417	Human ner	189	18	1.5	1579	7	ACD18055	Novel hum
117	18	1.5	544	5	ABX83757	Pepperm	190	18	1.5	1579	7	ACD18055	Novel hum
118	18	1.5	577	5	AAH87896	Pepperm	191	18	1.5	1579	7	ACD18055	Novel hum
119	18	1.5	669	7	ABX74583	Human CDN	192	18	1.5	1579	7	ACD18055	Novel hum
C 120	18	1.5	711	5	AA579439	DNA encod	193	18	1.5	1579	7	ACD18055	Novel hum
C 121	18	1.5	843	5	AA574361	DNA encod	194	18	1.5	1579	7	ACD18055	Novel hum
122	18	1.5	930	5	AA577162	DNA encod	195	18	1.5	1579	7	ACD18055	Novel hum
123	18	1.5	1060	3	AA477133	FtsZ homo	196	18	1.5	1579	7	ACD18055	Novel hum
124	18	1.5	1140	3	ADA70195	Rice gene	197	18	1.5	1579	7	ACD18055	Novel hum
125	18	1.5	1194	7	ABX70989	Novel hum	198	18	1.5	1579	7	ACD18055	Novel hum
126	18	1.5	1212	4	ABX11413	Drosophil	199	18	1.5	1579	7	ACD18055	Novel hum
C 127	18	1.5	1374	6	ABA03180	Human kid	200	18	1.5	1579	7	ACD18055	Novel hum
C 128	18	1.5	1374	6	ABT23506	Cellobich	201	18	1.5	1579	7	ACD18055	Novel hum
C 129	18	1.5	1438	3	AA11722	Human cel	202	18	1.5	1579	7	ACD18055	Novel hum
130	18	1.5	1438	3	AA11722	Human cel	203	18	1.5	1579	7	ACD18055	Novel hum
131	18	1.5	1438	9	AD60631	Human cel	204	18	1.5	1579	7	ACD18055	Novel hum
132	18	1.5	1438	6	ABQ99334	Human cod	205	18	1.5	1579	7	ACD18055	Novel hum
133	18	1.5	1465	6	ABA03179	Human kid	206	18	1.5	1579	7	ACD18055	Novel hum
134	18	1.5	1550	6	ABA03179	Human kid	207	18	1.5	1579	7	ACD18055	Novel hum
135	18	1.5	1579	3	AA596337	cDNA enco	208	18	1.5	1579	7	ACD18055	Novel hum
136	18	1.5	1579	4	AA546145	Human DNA	209	18	1.5	1579	7	ACD18055	Novel hum
137	18	1.5	1579	7	ABX78748	Human PRO	210	18	1.5	1579	7	ACD18055	Novel hum
138	18	1.5	1579	7	ACA75720	Novel hum	211	18	1.5	1579	7	ACD18055	Novel hum
139	18	1.5	1579	7	ACA71200	Human sec	212	18	1.5	1579	7	ACD18055	Novel hum
140	18	1.5	1579	7	ACC87728	Human sec	213	18	1.5	1579	7	ACD18055	Novel hum
141	18	1.5	1579	7	ACC87728	Human sec	214	18	1.5	1579	7	ACD18055	Novel hum
142	18	1.5	1579	7	ACD04287	Human sec	215	18	1.5	1579	7	ACD18055	Novel hum
143	18	1.5	1579	7	ACA9618	Human sec	216	18	1.5	1579	7	ACD18055	Novel hum
144	18	1.5	1579	7	ACA90463	Novel hum	217	18	1.5	1579	7	ACD18055	Novel hum
145	18	1.5	1579	7	ACC89570	Human sec	218	18	1.5	1579	7	ACD18055	Novel hum
146	18	1.5	1579	7	ACA98361	Novel hum	219	18	1.5	1579	7	ACD18055	Novel hum
147	18	1.5	1579	7	ACA94003	Human sec	220	18	1.5	1579	7	ACD18055	Novel hum
148	18	1.5	1579	7	ACA06131	cDNA enco	221	18	1.5	1579	7	ACD18055	Novel hum
149	18	1.5	1579	7	ACD15396	Human sec	222	18	1.5	1579	7	ACD18055	Novel hum
150	18	1.5	1579	7	ACD08983	Human sec	223	18	1.5	1579	7	ACD18055	Novel hum
151	18	1.5	1579	7	ACC96903	Human sec	224	18	1.5	1579	7	ACD18055	Novel hum
152	18	1.5	1579	7	ACF15624	Human sec	225	18	1.5	1579	7	ACD18055	Novel hum
153	18	1.5	1579	7	ACA72991	Human PRO	226	18	1.5	1579	7	ACD18055	Novel hum
154	18	1.5	1579	7	ACD03163	Novel hum	227	18	1.5	1579	7	ACD18055	Novel hum
155	18	1.5	1579	7	ACA92170	Novel hum	228	18	1.5	1579	7	ACD18055	Novel hum
156	18	1.5	1579	7	ACA89595	cdNA enco	229	18	1.5	1579	7	ACD18055	Novel hum
157	18	1.5	1579	7	ACA73605	Human sec	230	18	1.5	1579	7	ACD18055	Novel hum
158	18	1.5	1579	7	ACA05920	Human sec	231	18	1.5	1579	7	ACD18055	Novel hum
159	18	1.5	1579	7	ACA66754	cdNA enco	232	18	1.5	1579	7	ACD18055	Novel hum
160	18	1.5	1579	7	ACA66754	cdNA enco	233	18	1.5	1579	7	ACD18055	Novel hum
161	18	1.5	1579	7	ACF20329	Human sec	234	18	1.5	1579	7	ACD18055	Novel hum
162	18	1.5	1579	7	ACF19715	Human sec	235	18	1.5	1579	7	ACD18055	Novel hum
163	18	1.5	1579	7	ACD22003	Human sec	236	18	1.5	1579	7	ACD18055	Novel hum
164	18	1.5	1579	7	ACF13168	Human sec	237	18	1.5	1579	7	ACD18055	Novel hum
165	18	1.5	1579	7	ACD28800	Human sec	238	18	1.5	1579	7	ACD18055	Novel hum
166	18	1.5	1579	7	ACD5271	Human sec	239	18	1.5	1579	7	ACD18055	Novel hum
167	18	1.5	1579	7	ACF00320	Human sec	240	18	1.5	1579	7	ACD18055	Novel hum
168	18	1.5	1579	7	ACA72377	Novel hum	241	18	1.5	1579	7	ACD18055	Novel hum
169	18	1.5	1579	7	ACD04901	Novel hum	242	18	1.5	1579	7	ACD18055	Novel hum

243	18	1.5	1579	7	ACC92135
244	18	1.5	1579	7	ACA74992
245	18	1.5	1579	7	ACA91863
246	18	1.5	1579	7	ACA71507
247	18	1.5	1579	7	ACC90907
248	18	1.5	1579	7	ACA65917
249	18	1.5	1579	7	ACA95062
250	18	1.5	1579	7	ACD16624
251	18	1.5	1579	7	ACD15703
252	18	1.5	1579	7	ABX16806
253	18	1.5	1579	8	ACA97747
254	18	1.5	1579	8	ACA99196
255	18	1.5	1579	8	ACC91828
256	18	1.5	1579	8	ACA67897
257	18	1.5	1579	8	ACD11239
258	18	1.5	1579	8	ACD15089
259	18	1.5	1579	8	ACD11853
260	18	1.5	1579	8	ACC95982
261	18	1.5	1579	8	ACF16545
262	18	1.5	1579	8	ACF02653
263	18	1.5	1579	8	ACF02970
264	18	1.5	1579	8	ACF21557
265	18	1.5	1579	8	ACF10241
266	18	1.5	1579	8	ACF78134
267	18	1.5	1579	8	ACD46839
268	18	1.5	1579	8	ACD49602
269	18	1.5	1579	8	ADA76533
270	18	1.5	1579	8	ACF28369
271	18	1.5	1579	8	ACD89059
272	18	1.5	1579	8	ACD84454
273	18	1.5	1579	8	ACD99228
274	18	1.5	1579	8	ADA78193
275	18	1.5	1579	8	ACF48970
276	18	1.5	1579	8	ACD09290
277	18	1.5	1579	8	ACF12083
278	18	1.5	1579	8	ACF41317
279	18	1.5	1579	8	ACF15931
280	18	1.5	1579	8	ACF16238
281	18	1.5	1579	8	ACD32065
282	18	1.5	1579	8	ACF18873
283	18	1.5	1579	8	ACF09320
284	18	1.5	1579	8	ACF78441
285	18	1.5	1579	8	ACF52040
286	18	1.5	1579	8	ACF26527
287	18	1.5	1579	8	ACF24320
288	18	1.5	1579	8	ACF63431
289	18	1.5	1579	8	ACF50505
290	18	1.5	1579	8	ACH07976
291	18	1.5	1579	8	ACF13782
292	18	1.5	1579	8	ACD41708
293	18	1.5	1579	8	ACF32121
294	18	1.5	1579	8	ACF23399
295	18	1.5	1579	8	ACF40089
296	18	1.5	1579	8	ACD45611
297	18	1.5	1579	8	ACF53268
298	18	1.5	1579	8	ACF27448
299	18	1.5	1579	8	ACF45286
300	18	1.5	1579	8	ACF29904

ALIGNMENTS

RESULT 1	
ID	ABK61465 standard; cDNA; 1183 BP.
XX	ABK61465
AC	ABK61465;
XX	
DT	18-JUN-2002 (first entry)
DE	Human cDNA encoding protein NOV13.
XX	

Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis; cell signal processing disorder; metabolic pathway modulation disorder; diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; uterus cancer; immune response; graft-versus-host disease; acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease; hypertension; congenital heart defects; multiple sclerosis; inflammation; Albright hereditary osteodystrophy.

Homo sapiens.

WO200216599-A2.

28-FEB-2002.

27-AUG-2001; 2001WO-US026510.

25-AUG-2000; 2000US-0228191P.

08-FEB-2001; 2001US-0267300P.

20-FEB-2001; 2001US-0269961P.

20-MAR-2001; 2001US-0277337P.

(CURA-) CURAGEN CORP.

(CORT-) COR THERAPEUTICS INC.

Burgess CF, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA; Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R; WPI; 2002-280937/32.

P-PSDB; AAU91308.

New polypeptides for treating or preventing a disorder associated with them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.

Claim 1; Page 98; 263pp; English.

The invention relates to an isolated polypeptide (NOVX) a mature form of NOVX, a NOVX variant (differing by no more than 15%), the nucleotide encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14, 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it and antibody against it, are useful for treating or preventing (e.g. by gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy, atherosclerosis, a disorder related to cell signal processing and metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide and nucleic acids are also useful for determining the presence of predilection to the diseases. The NOVX nucleic acid and polypeptide are especially useful in therapeutic or prophylactic applications for disorders associated with aberrant NOVX expression or activity, e.g. cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus cancer), immune response, graft-versus-host disease, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension, congenital heart defects, multiple sclerosis, inflammation or Albright hereditary osteodystrophy and many other diseases listed in the specification. The DNA encoding the protein is useful in gene therapy for treating the conditions. This is also useful in detection assays, chromosome mapping, tissue typing, diagnostic or prognostic assays, or for developing a powerful assay system for functional analysis of various human disorders, as well as in diagnostic applications. The present sequence encodes a NOVX protein

Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1183;	DB 6;	Length 1183;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1183;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGCTAGAGCTCCAGGACCCACGCGCTGTCTCTGTGACAGAGCTCAAGGCGCCCTGGG	60	
Db	1	AGCTAGAGCTCCAGGACCCACGCGCTGTCTCTGTGACAGAGCTCAAGGCGCCCTGGG	60	
Qy	61	CTTTCCTCCCTCGCTCGGCTGTGCTTGGAGGAGTCCCGAGTCCAGATCCCTAGGAG	120	
Db	61	CTTTCCTCCCTCGCTCGGCTGTGCTTGGAGGAGTCCCGAGTCCAGATCCCTAGGAG	120	

121 CATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180
121 CATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180
181 ACCAAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAGTCTGGGTG 240
181 ACCAAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAGTCTGGGTG 240
241 TCTAGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTTAACCTGTCCA 300
241 TCTAGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTTAACCTGTCCA 300
301 GCCAGAGCATGGCTCTCAGCAGAGCTGTCTCCAGGCTTTGATGACAAACCAATTTCC 360
301 GCCAGAGCATGGCTCTCAGCAGAGCTGTCTCCAGGCTTTGATGACAAACCAATTTCC 360
361 CTCGATGATGCTTCTGAGTGTCTCTGAGGACCAATGGGAGTCTGCCAGCAGAG 420
361 CTCGATGATGCTTCTGAGTGTCTCTGAGGACCAATGGGAGTCTGCCAGCAGAG 420
421 AAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTCCAAAGCCAGGACCTGTGACCAT 480
421 AAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTCCAAAGCCAGGACCTGTGACCAT 480
481 GGNAGCAGAGAGAGAGGACAGCCAGCGCTGGCCCTGGGACGTTTCCCGGACAGTGGGCC 540
481 GGAAGCAGAGAGAGAGGACAGCCAGCGCTGGCCCTGGGACGTTTCCCGGACAGTGGGCC 540
541 GGCAGAGCTGTGCTCAGAGCTCGGGAGCCATTGACCATCTCTCTGAGGATGGAGACTG 600
541 GGCAGAGCTGTGCTCAGAGCTCGGGAGCCATTGACCATCTCTCTGAGGATGGAGACTG 600
601 GTGGACGGTGTCTGTAAGTCTCAGGAGAGAGTATAACATCCCGAGGTCACGTTGG 660
601 GTGGACGGTGTCTGTAAGTCTCAGGAGAGAGTATAACATCCCGAGGTCACGTTGG 660
661 CAAAGTCTCCATGGTGTGCTGATGAGGCTGATGAGGCTGAGCAGGAGGAAAGCAGAGGACTGCT 720
661 CAAAGTCTCCATGGTGTGCTGATGAGGCTGATGAGGCTGAGCAGGAGGAAAGCAGAGGACTGCT 720
721 GTTGTACTGGAAACCTGGAGGGGCTTCTCATTCGGGAGAGCAGACAGCAGAGAGG 780
721 GTTGTACTGGAAACCTGGAGGGGCTTCTCATTCGGGAGAGCAGACAGCAGAGAGG 780
781 CTCTTACTCTGCTGAGTGGCTGAGCGGCTGATCTGCGGACCGATCAGACACTA 840
781 CTCTTACTCTGCTGAGTGGCTGAGCGGCTGATCTGCGGACCGATCAGACACTA 840
841 CAGGATCCACTGCTTGGACAAATGGCTGGCTGTGATCATTCACCGCGCTCACCTTCCCTTC 900
841 CAGGATCCACTGCTTGGACAAATGGCTGGCTGTGATCATTCACCGCGCTCACCTTCCCTTC 900
901 ACTCAGGCGCTGGTGGACCAATTAATCTGAGCTGGGGATGATCTGCTGCTACTCTCA 960
901 ACTCAGGCGCTGGTGGACCAATTAATCTGAGCTGGGGATGATCTGCTGCTACTCTCA 960
961 GGAGCCCTGTGCTGAGAGGGCTGGCGGCTTCCCTGGCAAGGATATACCCCTACCTGT 1020
961 GGAGCCCTGTGCTGAGAGGGCTGGCGGCTTCCCTGGCAAGGATATACCCCTACCTGT 1020
1021 GACTGTGAGAGGACCACTAATCTGAAAGAGCTGGACGCTCCCTCTGTTTCTGA 1080
1021 GACTGTGAGAGGACCACTAATCTGAAAGAGCTGGACGCTCCCTCTGTTTCTGA 1080
1081 AGCTGCCACAGGGGAGGAGTCTTCTCAGTGGGCTCTCCGAGCTCCCTCAGCTTCTTA 1140
1081 AGCTGCCACAGGGGAGGAGTCTTCTCAGTGGGCTCTCCGAGCTCCCTCAGCTTCTTA 1140
1141 CATCAGCTGAATGACAGGCTGTCTTTGGATGATGCTAG 1183
1141 CATCAGCTGAATGACAGGCTGTCTTTGGATGATGCTAG 1183

RESULT 2
AAD43980
ID AAD43980 standard; cDNA; 2567 BP.
XX
AC AAD43980;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.
XX
KW Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;
KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;
KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;
KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;
KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;
KW multiple sclerosis; asthma; acute respiratory distress syndrome;
KW pulmonary disorder; dermatological; neuroprotective; gene; ss.
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 415..1200
FT /*tag= a
FT /product= "Human SLAP-2"
XX
XX WO200242457-A1.
XX
PD 30-MAY-2002.
XX
XX 20-NOV-2001; 2001WO-US043367.
XX
XX 22-NOV-2000; 2000US-0252545P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;
XX
XX WPI; 2002-463632/49.
XX
XX P-PSDB; AAE26357.
XX
XX Novel substantially purified human SH2/SH3-domain-containing adapter
XX polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic
XX intervention in immunological and inflammatory disorders and cancer.
XX
XX Claim 2; Fig 1; 85pp; English.
XX
XX The invention relates to a substantially purified human SH2/SH3-domain-
XX containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-
XX 2). The invention is useful for treating an immune disorder involving
XX hyperactivity of B- or T- lymphocytes in a mammal. The invention is
XX useful for screening for antagonists or inhibitors of the interaction of
XX hSLAP-2 with cellular signalling compounds, for diagnosing, treating or
XX preventing diseases or disorders associated with aberrant or uncontrolled
XX cellular signal transduction, for determining those cellular signalling
XX molecules which associate with hSLAP-2 and which provide critical signals
XX for cell activation, and as effectors in methods to affect T-cell
XX activation. The invention is useful in screening assays to identify and
XX detect candidate autoimmune diseases which may be caused by
XX potential use to treat autoimmune diseases which may be caused by
XX hyperactivated B cells, as well as to treat diseases which may be caused
XX by hyperactivated T cells, in addition to other immune system related
XX conditions, diseases, or disorders, T-cell and B-cell neoplasms,
XX inflammation disorders, diseases and conditions, rheumatoid arthritis,
XX osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's
XX and ulcerative colitis), allergies, particularly those involving
XX hyperactivity of B-cells and T-cells, or other immune cells, such as
XX mast cells or eosinophils, autoimmune diseases such as systemic lupus
XX erythematosus and multiple sclerosis, pulmonary diseases including
XX asthma, acute respiratory distress syndrome, and chronic obstructive
XX pulmonary disorder, tissue/organ rejection and cancer. The invention is
XX useful in gene therapy. The present sequence is human SLAP-2 cDNA
XX
XX Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Query Match 95.7%; Score 1132; DB 6; Length 2567;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCTAGAGCTCCAAGAGACCCACGCTGTCTGTGTGACAGAGCTCAAAGGGCCCTGGG 60
 Db 18 AGCTAGAGCTCCAAGAGACCCACGCTGTCTGTGTGACAGAGCTCAAAGGGCCCTGGG 77

QY 61 CTTTCCCTCCCTGGCTCGCTGTCTGTGGAGGGTCCCAAGTCCAGATCCCTAAGGAG 120
 Db 78 CTTTCCCTCCCTGGCTCGCTGTCTGTGGAGGGTCCCAAGTCCAGATCCCTAAGGAG 137

QY 121 CATGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180
 Db 138 CATGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 197

QY 181 ACCCAACCAACACCTAGGCTCTCCCTGAAGATCCCTCCAGGCTGAGAGAGTTCTGGGTG 240
 Db 198 ACCCAACCAACACCTAGGCTCTCCCTGAAGATCCCTCCAGGCTGAGAGAGTTCTGGGTG 257

QY 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTCCA 300
 Db 258 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTCCA 317

QY 301 GCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCC 360
 Db 318 GCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCC 377

QY 361 CTCGATGATGCTCTCTGAGTCTCTGCTGAGAACATGGGAAGTCTGCCAGCAGAG 420
 Db 378 CTCGATGATGCTCTCTGAGTCTCTGCTGAGAACATGGGAAGTCTGCCAGCAGAG 437

QY 421 AAAATCTCTGCCAAGCCCAAGCTTCTGAGTTCTCTGTCCAAAGCCAGGACCTGTGACCAT 480
 Db 438 AAAATCTCTGCCAAGCCCAAGCTTCTGAGTTCTCTGTCCAAAGCCAGGACCTGTGACCAT 497

QY 481 GGAAGCAGAGAGAGCAGAGCCACAGCCGTGGCCCTGGGCAATTTCCCGGAGAGTGGCCC 540
 Db 498 GGAAGCAGAGAGAGCAGAGCCACAGCCGTGGCCCTGGGCAATTTCCCGGAGAGTGGCCC 557

QY 541 GCCCAGAGTCTCGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGACTG 600
 Db 558 GCCCAGAGTCTCGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGACTG 617

QY 601 GTGGACGCTGTCTGTCTGAGCTCAGGACAGAGATTAACATCCCCAGCGTCCACGTGG 660
 Db 618 GTGGACGCTGTCTGTCTGAGCTCAGGACAGAGATTAACATCCCCAGCGTCCACGTGG 677

QY 661 CAAAGTCTCCATGGTGGCTGTATGAGGCTGTAGGCGCTGAGCGGAGAAAGCAGAGGACTGCT 720
 Db 678 CAAAGTCTCCATGGTGGCTGTATGAGGCTGTAGGCGCTGAGCGGAGAAAGCAGAGGACTGCT 737

QY 721 GTTGTATCCTGGGAACCCCTGGAGGGGCTTCTCTCATCCGGGAGAGCCAGACAGAGAGG 780
 Db 738 GTTGTATCCTGGGAACCCCTGGAGGGGCTTCTCTCATCCGGGAGAGCCAGACAGAGAGG 797

QY 781 CTCCTACTCTGTCTGAGTCCGCTCAGCGCCCTCGATCCTGGGACCGGATCAGACATA 840
 Db 798 CTCCTACTCTGTCTGAGTCCGCTCAGCGCCCTCGATCCTGGGACCGGATCAGACATA 857

QY 841 CAGGATCCACTCCCTTGACATGGCTGGCTGTACATCTCACCGGCGCTCACCTTCCCTC 900
 Db 858 CAGGATCCACTCCCTTGACATGGCTGGCTGTACATCTCACCGGCGCTCACCTTCCCTC 917

QY 901 ACTCAGGCGCTGGTGGACCATTACTCTGAGTGGCGGATGACATCTCTGCTACTCAA 960
 Db 918 ACTCAGGCGCTGGTGGACCATTACTCTGAGTGGCGGATGACATCTCTGCTACTCAA 977

QY 961 GGAGCCCTGTGCTGAGAGGGCTGGCCCGCTCCCTGCGAGGATATACCCCTACCTGT 1020
 Db 978 GGAGCCCTGTGCTGAGAGGGCTGGCCCGCTCCCTGCGAGGATATACCCCTACCTGT 1037

QY 1021 GACTGTGACAGAGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGA 1080
 Db 1038 GACTGTGACAGAGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGA 1097

QY 1081 AGCTGCCACAGGGGAGAGTCTCTTCTCAGTGAAGGTTCTCCGGAGTCCCTCAGCTTCTA 1140
 Db 1098 AGCTGCCACAGGGGAGAGTCTCTTCTCAGTGAAGGTTCTCCGGAGTCCCTCAGCTTCTA 1157

QY 1141 CATCAGCTCAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183
 Db 1158 CATCAGCTCAATGACAGGCTGTCTCTTTGGATGATGCTAG 1200

RESULT 3

AAI44089

ID AAI44089 standard; cDNA; 786 BP.

XX AAI44089;

AC AAI44089;

DT 03-OCT-2002 (first entry)

XX Human modulator of antigen receptor signalling protein coding sequence.
 XX Human; gene; ss; gene therapy; modulator of antigen receptor signalling;
 KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
 KW immunosuppression; myeloproliferative disorder; breast cancer.

OS Homo sapiens.

XX Key Location/Qualifiers
 CDS 1..786
 FT /*tag= a
 FT /product= "Human MARS protein"

PN WO200242452-A2.

PD 30-MAY-2002.

PF 26-NOV-2001; 2001WO-CA001662.

PR 27-NOV-2000; 2000CA-02324663.

XX (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX Meglade JC, Loreto MP;

XX WPI; 2002-566564/60.

XX P-PSDB; AAO15457.

XX New isolated modulator of antigen receptor signaling protein or its
 fragment, useful for treating malignant disorders such as myeloid
 malignancies, autoimmune disorders and myeloproliferative disorders.

XX Claim 12; Page 75; 110pp; English.

XX The invention comprises the amino acid and coding sequences of modulator
 of antigen receptor signalling (MARS) proteins. The MARS protein is a
 putative tumour suppressor gene and exhibits structural and sequence
 similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
 protein sequences of the invention are useful for the treatment of
 myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
 disorders, immunosuppression, myeloproliferative disorders and
 malignancies related to the de-regulation of tyrosine kinases (e.g.
 breast cancer). The present cDNA sequence encodes a human MARS protein

XX Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

XX Query Match 62.1%; Score 735; DB 6; Length 786;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGGAAGTCTGCCAGCAGAGAGAAAATCTGCCAAGCCCAAGCTTGTCTCTGTC 457

Db 1 ATGGGAGTCTGCCAGCAGAGAAATCTTCCAGCCCAAGCTTGAGTTCTCTGTC 60
 QY 458 CAAGCCAGGACCTGTGACCAATGGAAGCAGAGAGAGGACCAAGCCGTCCTG 517
 Db 61 CAAGCCAGGACCTGTGACCAATGGAAGCAGAGAGAGGACCAAGCCGTCCTG 120
 QY 518 GCGAGTTTCCCGCAGGTTGGCCGCGAGCTGCTGAGACTCGGGAGCCATTGACC 577
 Db 121 GCGAGTTTCCCGCAGGTTGGCCGCGAGCTGCTGAGACTCGGGAGCCATTGACC 180
 QY 578 ATCGTCTCTGAGGATGAGACTGCTGGAGCGTGTCTGAGTCTCAGGAGAGAT 637
 Db 181 ATCGTCTCTGAGGATGAGACTGCTGGAGCGTGTCTGAGTCTCAGGAGAGAT 240
 QY 638 AACATCCCGAGCTGACCTGGGCAAGTCTCCATGGGTGGTGTATGAGGCTGAGC 697
 Db 241 AACATCCCGAGCTGACCTGGGCAAGTCTCCATGGGTGGTGTATGAGGCTGAGC 300
 QY 698 AGGAGAAAGCAGAGAACTGCTGTTTACCTGGACCTCGAGGGGCTTCTCATC 757
 Db 301 AGGAGAAAGCAGAGAACTGCTGTTTACCTGGACCTCGAGGGGCTTCTCATC 360
 QY 758 CCGGAGAGCCAGACAGGAGAGGCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGCA 817
 Db 361 CCGGAGAGCCAGACAGGAGAGGCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGCA 420
 QY 818 TCTGGGCGGATCAGACACTACAGGATCCACTGCTTGAATGGCTGGCTGACATC 877
 Db 421 TCTGGGCGGATCAGACACTACAGGATCCACTGCTTGAATGGCTGGCTGACATC 480
 QY 878 TCACCGCGCTCACCCTTCCCTCACTCCAGGCGCTGTGGACCACTTACTCTGAGTGGC 937
 Db 481 TCACCGCGCTCACCCTTCCCTCACTCCAGGCGCTGTGGACCACTTACTCTGAGTGGC 540
 QY 938 GATGACATCTGCTGCTTACTCAAGGAGCCCTGTGCTGAGAGGCTGGCCGCTCCCT 997
 Db 541 GATGACATCTGCTGCTTACTCAAGGAGCCCTGTGCTGAGAGGCTGGCCGCTCCCT 600
 QY 998 GGCAGGATATACCCCTACTGCTGACTGTGACAGAGACCACTCACTGGAAGAGGTG 1057
 Db 601 GGCAGGATATACCCCTACTGCTGACTGTGACAGAGACCACTCACTGGAAGAGGTG 660
 QY 1058 GACAGTCCCTCTGTTTCTGAAGCTGCCAGGGAGAGTCTCTCTCAGTGAAGGT 1117
 Db 661 GACAGTCCCTCTGTTTCTGAAGCTGCCAGGGAGAGTCTCTCTCAGTGAAGGT 720
 QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGAGGCTGTCTTTGATGAT 1177
 Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGAGGCTGTCTTTGATGAT 780
 QY 1178 GCGTAG 1183
 Db 781 GCGTAG 786

RESULT 4
 ABQ74343
 ID ABQ74343 standard; cDNA; 786 BP.

XX ABQ74343;
 XX

DT 15-OCT-2002 (first entry)

XX Human Src-like inhibitory molecule (SLIM) encoding cDNA.

XX Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;
 KW SLAP; inhibitor; anti-inflammatory; immunosuppressive; anti-HIV;
 KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;
 KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;
 KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;
 KW gene; ss.

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 1.786
 FT /tag= a
 FT /product= "SLIM"
 FT /note= "Src-like inhibitory molecule"
 XX W0200255707-A2.
 EN 18-JUL-2002.
 PD 10-JAN-2002; 2002MO-US000718.
 PF 10-JAN-2001; 2001US-0260953P.
 PR (RIGE-) RIGEL PHARM INC.
 PA Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y,
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;
 XX WPI: 2002-575432/61.
 DR P-PSDB; ABP52187.
 DR
 XX New src-like inhibitory molecule protein, useful for treating
 PT immunodeficiency disorders and inflammatory disorders, comprises N-
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.
 XX Claim 3; Fig 2A; 91pp; English.
 CC The present sequence encodes the human Src-like inhibitory molecule
 CC (SLIM) protein (I). The present invention describes a SLIM protein
 CC comprising an N-terminal myristylation sequence, an N-terminal SH2
 CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising
 CC an N-terminal myristylation sequence and an N-terminal SH2 domain which
 CC is unable to bind to Cbl. (I) has antinflammatory, immunosuppressive and
 CC anti-HIV activities, and can be used as a modulator of lymphocyte
 CC activation, and of ubiquitination of a Cbl target protein, and in gene
 CC therapy. (I) is useful for screening a bioactive agent capable of binding
 CC to SLIM. (I) is also useful for screening a bioactive agent capable of
 CC modulating SLIM binding. (I) or its fragments is useful in the study or
 CC in the treatment of conditions which involves this function or
 CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or
 CC prevent SLIM associated disorders. (II) or the polynucleotide encoding it
 CC (II) is useful for modulating leukocyte and/or platelet activation, for
 CC modulating antigen receptor-induced signaling and activation in
 CC leukocyte and/or platelets and for modulating antigen receptor-induced
 CC signaling and activation in lymphocytes and/or mast cells. (I) or (II)
 CC is also useful for modulating the basal activity of lymphocytes. (I) or
 CC (II) is useful in the treatment of immunodeficiency disorders, such as
 CC acquired immunodeficiency syndrome (AIDS), for the prevention and
 CC treatment of acute inflammatory disorders, chronic inflammatory
 CC disorders, autoimmune disorder and transplant rejection
 XX Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;
 SQ

Query Match 62.1%; Score 735; DB 6; Length 786;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTTCCAGCCCAAGCTTGAGTTCTCTGTC 457

Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTTCCAGCCCAAGCTTGAGTTCTCTGTC 60

QY 458 CAAGCCAGGACCTGTGACCAATGGAAGCAGAGAGAGGACCAAGCCGTCCTG 517

Db 61 CAAGCCAGGACCTGTGACCAATGGAAGCAGAGAGAGGACCAAGCCGTCCTG 120

QY 518 GCGAGTTTCCCGCAGGTTGGCCGCGAGCTGCTGAGACTCGGGAGCCATTGACC 577

Db 121 GCGAGTTTCCCGCAGGTTGGCCGCGAGCTGCTGAGACTCGGGAGCCATTGACC 180

QY 578 ATCGTCTCTGAGGATGAGACTGCTGGACCGTGTCTCTGAGTCTCAGGAGAGTAT 637

```
Db 191 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGTGAAGTCTCAGGACAGAGATAT 240
Qy 638 AACATCCCGACGCTCCAGTGGGCAAACTCTCCCATGGTGGCTGTATGAGGGGCTCAGC 697
Db 241 AACATCCCGACGCTCCAGTGGGCAAACTCTCCCATGGTGGCTGTATGAGGGGCTCAGC 300
Qy 698 AGGAGGAAGCAGAGGAAGTCTGTGTACTGGGAACCTGGAGGGGCTTCTCTCATC 757
Db 301 AGGAGGAAGCAGAGGAAGTCTGTGTACTGGGAACCTGGAGGGGCTTCTCTCATC 360
Qy 758 CGGAGAGCCAGACACAGAGAGGCTCTTACTCTCTGTGAGTCCGCTCAGCGCCCTGCA 817
Db 361 CGGAGAGCCAGACACAGAGAGGCTCTTACTCTCTGTGAGTCCGCTCAGCGCCCTGCA 420
Qy 818 TCCTGGGACCGATCAGACACTACAGGATCACTGCTTGAATGGCTGCTGTACATC 877
Db 421 TCCTGGGACCGATCAGACACTACAGGATCACTGCTTGAATGGCTGCTGTACATC 480
Qy 878 TCACCGCGCTCAGCTTCCCTCACTCCAGGCGCTGTGGACCACTTACTGTAGCTGGCG 937
Db 481 TCACCGCGCTCAGCTTCCCTCACTCCAGGCGCTGTGGACCACTTACTGTAGCTGGCG 540
Qy 938 GATGACATCTGCTCTACTCAAGAGCCCTGTGTCTGCAAGGGCTGGCCGCTCCT 997
Db 541 GATGACATCTGCTCTACTCAAGAGCCCTGTGTCTGCAAGGGCTGGCCGCTCCT 600
Qy 998 GCGAGGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCAACTGGAAGAGCTG 1057
Db 601 GCGAGGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCAACTGGAAGAGCTG 660
Qy 1058 GACAGCTCCCTCTCTTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGGT 1117
Db 661 GACAGCTCCCTCTCTTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGGT 720
Qy 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGAAATGACGAGGCTGTCTCTTCTGATGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGAAATGACGAGGCTGTCTCTTCTGATGAT 780
Qy 1178 GCCTAG 1183
Db 781 GCCTAG 786

RESULT 5
ABQ98670
ID ABQ98670 standard; DNA; 763 BP.
XX
AC ABQ98670;
XX
XX 04-NOV-2002 (first entry)
XX
DE Human ORF477 coding sequence.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;
KW Antinflammatory; Gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
XX US2002082206-A1.
XX
XX 27-JUN-2002.
XX
XX 30-MAY-2001; 2001US-00867550.
XX
XX 30-MAY-2000; 2000US-0208427P.
XX
XX (LEAC/) LEACH M D.
XX
XX (MEHR/) MEHRABAN F.
XX
XX (CONL/) CONLEY P B.
```

PA (TOPP/) TOPPER J N.
XX (LAWD/) LAW D.

Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

WPI; 2002-626554/67.

P-PSDB; ABP64107.

New polypeptide designated ORFX are present in human atherogenic cells and are useful to prevent and treat ORFX-associated disorders including cancer, allergy, wound healing or autoimmune, cardiovascular or inflammatory disease.

Claim 2; SEQ ID NO 953; 78pp; English.

The present invention relates to novel human ORFX polypeptides and their coding sequences (ABP63631-ABP64681 and ABQ989267). The sequences were discovered in human atherogenic cells, in particular in platelets and human umbilical vein endothelial cells (HUVEC) and are expressed in many other tissues as well. Atherogenic cells are cells which have the potential to develop atherosclerotic plaques. The ORFX polypeptides and nucleic acids are useful for treating or preventing a pathological condition associated with an ORFX-associated disorder, e.g. cancer, cardiovascular disease, allergy, autoimmune disease, wound healing, blood coagulation disorders or inflammatory disorders. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov/sequence.html?DocID=2002082206

Seq Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 55.5%; Score 657; DB 6; Length 763;
Best Local Similarity 99.7%; Pred. No. 5.6e-300;
Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 117 GGAGCATGGGCGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTG 176
Db 5 GGAGCATGGGCGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTG 64
Qy 177 AGCTACCCAAACCAACACACCTAGCTCTCCCTGAAGATCTCTCCAGGCTGAGAGATTCTG 236
Db 65 AGCTACCCAAACCAACACCTAGCTCTCCCTGAAGATCTCTCCAGGCTGAGAGATTCTG 124
Qy 237 GGTGTCTTAGGACCAAGGACACTGGGACACTTCCAGAGGGCCCCCAAGCCCTTACCTG 296
Db 125 GATGTCTTAGGACCAAGGACACTGGGACACTTCCAGAGGGCCCCCAAGCCCTTACCTG 184
Qy 297 TCCAGCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCAGAGCCCTTTCATGACAAACCAAT 356
Db 185 TCCAGCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCAGAGCCCTTTCATGACAAACCAAT 244
Qy 357 TTCCCTCGATGATGTGCTTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCA 416
Db 245 TTCCCTCGATGATGTGCTTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCCAGCA 304
Qy 417 GAGAAATCTCTGCCAAGCCCCAAGCTTGAGTTCCTCTGTCAGAGGCCAGGGACCTGTGA 476
Db 305 GAGAAATCTCTGCCAAGCCCCAAGCTTGAGTTCCTCTGTCAGAGGCCAGGGACCTGTGA 364
Qy 477 CCATGGAAGCAGAGAGAGCAAGGCCACAGCCCTGGGCCCTGGGCAAGTTCCTGGGCAAGTG 536
Db 365 CCATGGAAGCAGAGAGAGCAAGGCCACAGCCCTGGGCCCTGGGCAAGTTCCTGGGCAAGTG 424
Qy 537 GCGCGCCGAGCTGTGCTGAGTGTGCTGAGGAGCCATTCATGACCTGCTCTGAGGATGGAG 596
Db 425 GCGCGCCGAGCTGTGCTGAGTGTGCTGAGGAGCCATTCATGACCTGCTCTGAGGATGGAG 484
Qy 597 ACTGGTGGACGGTGTCTGTCTGAAAGTCTCAGGACAGAGATATAACATCCCCAGGGTCCACG 656
Db 485 ACTGGTGGACGGTGTCTGTCTGAAAGTCTCAGGACAGAGATATAACATCCCCAGGGTCCACG 544
Qy 657 TGGGCAAGTCTCCATGGGTGGTGTATGAGGCTGTGATGAGGCTGTGACGAGGAGAAAGCAGAGAAC 716

Db 545 TGCCCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGCAGGAGAGAGAGAGAAC 604
 Qy 717 TGCTGTTGTTACTGTGGAAACCTTGAGGGGCTTCTCTATCCGGAGAGCCAGACCA 776
 Db 605 TGCTGTTGTTACTGTGGAAACCTTGAGGGGCTTCTCTATCCGGAGAGCCAGACCA 664
 Qy 777 GAGGCTCTTACTCTGTGTCAGTCCGCTCAGCGGCTGTCATCTGGGACCGGATCAGAC 836
 Db 665 GAGGCTCTTACTCTGTGTCAGTCCGCTCAGCGGCTGTCATCTGGGACCGGATCAGAC 724
 Qy 837 ACTACAGGATCCATGCTGCTGACCAATGCGTGTGCTGTACA 875
 Db 725 ACTACAGGATCCATGCTGCTGACCAATGCGTGTGCTGTACA 763

RESULT 6

AAC77202
 ID AAC77202 standard; cDNA; 837 BP.
 XX
 AC AAC77202;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2757 polynucleotide sequence SEQ ID NO:5513.
 XX
 KW Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US008621.
 XX
 PR 31-MAR-1999; 99US-0127607P.
 PR 02-APR-1999; 99US-0127636P.
 PR 05-APR-1999; 99US-0127728P.
 PR 30-MAR-2000; 2000US-00540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimketa RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR P-PSDB; AAB42993.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX
 PS Claim 5; Page 4692-4693; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosolic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; neurotropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;

CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 SQ Sequence 837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;
 Query Match 50.0%; Score 592; DB 3; Length 837;
 Best Local Similarity 99.6%; Pred. No. 2.8e-269;
 Matches 742; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 439 AAGCTTGAGTTCCTCTGTCCAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAA 498
 Db 3 AAGCTTGAGTTCCTCTGTCCAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAA 62
 Qy 499 GGGCACAGCGTGGCCCTGGGCAAGTTTCCGGCAGGTGGCCGGCCGCGAGTGTGCTGAG 558
 Db 63 GGGCACAGCGCGCGCTGGGCAAGTTTCCGGCAGGTGGCCGGCCGCGAGTGTGCTGAG 122
 Qy 559 ACTCGGGGCGCCATTGACCATCTCTCTGAGGATGAGACTGTGTGACCGTGTCTCTCA 618
 Db 123 ACTCGGGGCGCCATTGACCATCTCTCTGAGGATGAGACTGTGTGACCGTGTCTCTCA 182
 Qy 619 AGTCTCAGGCAGAGAGTATAAATCCCGACGCTCCACGTGGGCAAGTCTCCCATCGGTG 678
 Db 183 AGTCTCAGGCAGAGAGTATAAATCCCGACGCTCCACGTGGGCAAGTCTCCCATCGGTG 242
 Qy 679 GCTGTATGAGGGCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 738
 Db 243 GCTGTATGAGGGCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 302
 Qy 739 TGGAGGGGCGCTTCTCTCATCCGGAGAGCCAGACCCAGGAGGAGGAGGAGGAGGAGGAG 798
 Db 303 TGGAGGGGCGCTTCTCTCATCCGGAGAGCCAGACCCAGGAGGAGGAGGAGGAGGAGGAG 362
 Qy 799 CCGCTCAGCGCGCTGCTGATCTCAGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 858
 Db 363 CCGCTCAGCGCGCTGCTGATCTCAGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 422
 Qy 859 CAATGGCTGGCTGTACATCTCAGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 918
 Db 423 CAATGGCTGGCTGTACATCTCAGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 Qy 919 CCATTACTGTAGCTGGCGGATGACATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 978
 Db 483 CCATTACTGTAGCTGGCGGATGACATCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 542
 Qy 979 GAGGGCTGGCGCGCTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1038
 Db 543 GAGGGCTGGCGCGCTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602
 Qy 1039 ACTCAACTGGAAGAGCTGGAGAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1098
 Db 603 ACTCAACTGGAAGAGCTGGAGAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 662
 Qy 1099 GTCTCTCTCAGTGGAGGCTTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1158
 Db 663 GTCTCTCTCAGTGGAGGCTTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
 Qy 1159 GGCTGTCTCTTTGGATGATGCTCTAG 1183
 Db 723 GGCTGTCTCTTTGGATGATGCTCTAG 747

Db 181 ATGCTCTCTGAGGATGGAGCTGGTGGACGGTGTCTGTCTGAAAGTCTTCAGCAGAGAGTAT 240
Qy 638 AACATCCCGAGCGTCCACGTGGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 697
Db 241 AACATCCCGAGCGTCCACGTGGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300
Qy 698 AGGAG 757
Db 301 AGGAG 360
Qy 758 CGGAG 817
Db 361 CGGAG 420
Qy 818 TCCTGGGACCGGATCAGACACTACAGACTACAGACTACAGACTACAGACTACAGACTACAG 877
Db 421 TCCTGGGACCGGATCAGACACTACAGACTACAGACTACAGACTACAGACTACAGACTACAG 480
Qy 878 TCACCGGCGCTCACCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCG 931
Db 481 TCACCGGCGCTCACCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCG 534

RESULT 8
ABO99374
ID ABO99374 standard; cDNA; 1413 BP.
XX AC ABO99374;
XX DT 25-FEB-2003 (first entry)
XX DE Human coding sequence SEQ ID 107.
XX KW Human; expressed sequence tag; EST; chromosome 20;
KW haematopoietic disorder; central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
KW antileukemic; antineoplastic; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX OS Homo sapiens.
XX PN WO200259260-A2.
XX PD 01-AUG-2002.
XX PF 16-NOV-2001; 2001WO-US042950.
XX PR 17-NOV-2000; 2000US-00714936.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX DR WPI; 2002-590824/63.
XX DR N-PSDB; ABP64788.
XX PT New isolated polynucleotide, useful in research, diagnostic or
XX therapeutic methods, e.g. preventing or treating disorders involving
XX aberrant protein expression or biological activity.
XX PS Claim 1; SEQ ID NO 107; 394pp; English.
XX CC The present invention relates to novel human coding sequences (ABO99268-
XX ABO99608) and proteins (ABP6482-ABP6502). The sequences are useful in
XX therapeutic, diagnostic and research methods. The polynucleotides may be
XX used in the field of molecular biology as hybridisation probes, primers
XX for PCR, for chromosome and gene mapping, for the recombinant production
XX of protein, or in generation of anti-sense DNA or RNA. The
XX polynucleotides are useful in diagnostics as expressed sequence tags

AAAL44090
ID AAL44090 standard; cDNA; 737 BP.
XX AC AAL44090;
XX DT 03-OCT-2002 (first entry)
XX DE Mouse MARS short isoform protein coding sequence.
XX KW Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer.
XX OS Mus sp.
XX FH Key Location/Qualifiers
XX CDS 1..633
XX FT /*tag= a
XX FT /product= "Mouse MARS short isoform protein"
XX FT
XX PN WO200242452-A2.
XX PD 30-MAY-2002.
XX PF 26-NOV-2001; 2001WO-CA001662.
XX PR 27-NOV-2000; 2000CA-02324663.
XX PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX PI Mcglade JC, Loreto MP;
XX DR WPI; 2002-566564/60.
XX DR P-PSDB; AMO15458.
XX KW New isolated modulator of antigen receptor signaling protein or its
KW fragment, useful for treating malignant disorders such as myeloid
KW malignancies, autoimmune disorders and myeloproliferative disorders.
XX PS Claim 9; Page 77; 110pp; English.
XX CC The invention comprises the amino acid and coding sequences of modulator
XX of antigen receptor signalling (MARS) proteins. The MARS protein is a
XX putative tumour suppressor gene and exhibits structural and sequence
XX similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
XX protein sequences of the invention are useful for the treatment of
XX myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
XX disorders, immunosuppression, myeloproliferative disorders and
XX malignancies related to the de-regulation of tyrosine kinases (e.g.
XX breast cancer). The present cDNA sequence encodes a mouse MARS protein
XX
XX SQ Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;
Query Match 40.8%; Score 483; DB 6; Length 737;
Best Local Similarity 99.8%; Pred. No. 8.9e-218;
Matches 533; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 398 ATGGAGTCTCCAGCAG 457
Db 1 ATGGAGTCTCCAGCAG 60
Qy 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 517
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy 518 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTGCTGAGACTGGGGAGCCATTGACC 577
Db 121 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTGCTGAGACTGGGGAGCCATTGACC 180
Qy 578 ATCTCTCTGAGATGGAGACTGGTGGAGGCTGTCTGGAAGTCTTCAGGAGAGAGATAT 637

CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotential or pluripotential state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;

Query Match 39.6%; Score 468; DB 6; Length 1413;
 Best Local Similarity 99.6%; Pred. No. 1e-210;
 Matches 568; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 345 TGCAACCAATTCCTCGATGATGCTCTGAGTCTGCTGAGAACATGGGAA 404
 DB 1 TGCAACCAATTCCTCGATGATGCTCTGAGTCTGCTGAGAACATGGGAG 60
 QY 405 GTCTGCCAGCAGAGAAATCTGCCAAGCCAGCTTGATTCCTCTGCCAAGCC 464
 DB 61 GTCTGCCAGCAGAGAAATCTGCCAAGCCAGCTTGATTCCTCTGCCAAGCC 120
 QY 465 AGGACCTGTGACCATGGAAGCAGAGAGAGCAAGCCACAGCCCTGCGCCATG 524
 DB 121 AGGACCTGTGACCATGGAAGCAGAGAGAGCAAGCCACAGCCCTGCGCCATG 180
 QY 525 TCCCGCAGTGGCCGCGGAGCTGTCTGAGACTCGGGAGCCATTGACCATCGTCT 584
 DB 181 TCCCGCAGTGGCCGCGGAGCTGTCTGAGACTCGGGAGCCATTGACCATCGTCT 240
 QY 585 CTGAGATGAGACATGCTGACCGTGTCTGAAAGTCTCAGGCAGAGAGTAAACATCC 644
 DB 241 CTGAGATGAGACATGCTGACCGTGTCTGAAAGTCTCAGGCAGAGAGTAAACATCC 300
 QY 645 CCAGGCTCCAGTGGCAGAGTCTCCATGGGTGCTGTATGAGGCTGAGCAGGAGA 704
 DB 301 CCAGGCTCCAGTGGCAGAGTCTCCATGGGTGCTGTATGAGGCTGAGCAGGAGA 360
 QY 705 AAGCAGAGAACTGCTGTGTTTACCTGGGAAACCTTGAGGGGCTTCTCTATCCGGGAGA 764
 DB 361 AAGCAGAGAACTGCTGTGTTTACCTGGGAAACCTTGAGGGGCTTCTCTATCCGGGAGA 420
 QY 765 GCCAGACAGGAGAGGCTTCTACTCTGTAGTCCGCTGACCGCCCTGCATCTCGG 824
 DB 421 GCCAGACAGGAGAGGCTTCTACTCTGTAGTCCGCTGACCGCCCTGCATCTCGG 480
 QY 825 ACCGATCAGACATACAGGATCCACTGCTTGACAAATGGCTGGCTGTACATCTCACCGC 884
 DB 481 ACCGATCAGACATACAGGATCCACTGCTTGACAAATGGCTGGCTGTACATCTCACCGC 540
 QY 885 GCTTCACTTCCCTCACTCCAGGCGCTGG 914
 DB 541 GCTTCACTTCCCTCACTCCAGGCGCTGG 570

RESULT 9
 AAS74748
 ID AAS74748 standard; cDNA; 603 BP.
 XX
 AC AAS74748;
 XX

DT 13-FEB-2002 (first entry)
 XX DNA encoding novel human diagnostic protein #10552.
 XX DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG10561.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 10552; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;
 Query Match 34.2%; Score 405; DB 5; Length 603;
 Best Local Similarity 100.0%; Pred. No. 6.3e-181;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 779 GGCTCTTACTCTGTGTCAGTCCGCTCAGCGCCCTGTCATCTGGACCGGATCAGAC 838
 DB 199 GGCTCTTACTCTGTGTCAGTCCGCTCAGCGCCCTGTCATCTGGACCGGATCAGAC 258
 QY 839 TACAGGATCCATGCTCTTGCATATGCTGCTGTATCTCACCAGGCTCACCCTTCCC 898
 DB 259 TACAGGATCCATGCTCTTGCATATGCTGCTGTATCTCACCAGGCTCACCCTTCCC 318
 QY 899 TCACTCCAGGCTCTGTGGACCAATCTCTGAGTGGCGGATGACATCTGCTGCTACTC 958
 DB 319 TCACTCCAGGCTCTGTGGACCAATCTCTGAGTGGCGGATGACATCTGCTGCTACTC 378
 QY 959 AAGGAGCCCTGTGTCTCTGAGAGGGTGGCCGCTCCCTGGCAAGATATACCCCTACCT 1018

XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX WPI: 2002-626554/67.
DR P-PSDB; ABP64106.
XX
XX New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
XX Claim 2; SEQ ID NO 951; 78pp; English.
XX
XX The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=2002082206
XX
XX Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;
SQ
Query Match 29.4%; Score 348; DB 6; Length 444;
Best Local Similarity 100.0%; Pred. No. 5.4e-154;
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 242 CCTAGACCAAGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTACCTGTCCAG 301
Db 1 CCTAGACCAAGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTACCTGTCCAG 60
QY 302 CCAGAGCTGTCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCC 361
Db 61 CCAGAGCTGTCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCC 120
QY 362 TCGATGATGTCTCTGAGTGTCTCTGAGCAACTGGCAAGTTCGCCAGCAGAGA 421
Db 121 TCGATGATGTCTCTGAGTGTCTCTGAGCAACTGGCAAGTTCGCCAGCAGAGA 180
QY 422 AAATCTTGCCAAAGCCCAAGCTTGAGTTCCTGTGTCCAAAGCCAGGACCTGTGACCATG 481
Db 181 AAATCTTGCCAAAGCCCAAGCTTGAGTTCCTGTGTCCAAAGCCAGGACCTGTGACCATG 240
QY 482 GAAGCAGAGAGAGCAAGCCCAAGCCCTGAGCCCTGGCCGATTTCCCGGAGGTGGCCG 541
Db 241 GAAGCAGAGAGAGCAAGCCCAAGCCCTGAGCCCTGGCCGATTTCCCGGAGGTGGCCG 300
QY 542 GCCAGCTGTCTGAGTGTGGGAGCCATTTGACCATGCTCTCTGAG 589
Db 301 GCCAGCTGTCTGAGTGTGGGAGCCATTTGACCATGCTCTCTGAG 348
RESULT 12
ABQ99151
ID ABQ99151 standard; DNA; 875 BP.
XX
XX AC ABQ99151;
XX
XX DT 04-NOV-2002 (first entry)
XX
XX DE Human ORF958 coding sequence.
XX
XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;
KW Anti-inflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.

XX Homo sapiens.
XX OS
XX US2002082206-A1.
XX
XX 27-JUN-2002.
XX
XX 30-MAY-2001; 2001US-00867550.
XX
XX 30-MAY-2000; 2000US-0208427P.
XX
XX (LEACH/) LEACH M D.
XX PA (MEHR/) MEHRABAN F.
XX PA (CONL/) CONLEY P B.
XX PA (TOPP/) TOPPER J N.
XX PA (LAWD/) LAW D.
XX
XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
XX WPI: 2002-626554/67.
XX P-PSDB; ABP64588.
XX
XX New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
XX Claim 2; SEQ ID NO 1915; 78pp; English.
XX
XX The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=2002082206
XX
XX Sequence 875 BP; 205 A; 259 C; 225 G; 185 T; 0 U; 1 Other;
SQ
Query Match 28.8%; Score 341; DB 6; Length 875;
Best Local Similarity 100.0%; Pred. No. 1.1e-150;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 843 GGATCCACTGCCTTGACAAATGGCTGGCTGTACATCTCACCGCCCTACCTTCCCTCAC 902
Db 2 GGATCCACTGCCTTGACAAATGGCTGGCTGTACATCTCACCGCCCTACCTTCCCTCAC 61
QY 903 TCCAGGCCCTGGTGGACCAATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTCAAG 962
Db 62 TCCAGGCCCTGGTGGACCAATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTCAAG 121
QY 963 AGCCCTGTGTCTGCGAGAGGGCTGGCCCGCTCCCTCGCAAGGATATACCCCTACCTGTGA 1022
Db 122 AGCCCTGTGTCTGCGAGAGGGCTGGCCCGCTCCCTCGCAAGGATATACCCCTACCTGTGA 181
QY 1023 CTGTGCGAGGACACCACTCAACTGGAAGAGCTGGAGCTCCCTCTGTTTCTGAAG 1082
Db 182 CTGTGCGAGGACACCACTCAACTGGAAGAGCTGGAGCTCCCTCTGTTTCTGAAG 241
QY 1083 CTGCGACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACA 1142
Db 242 CTGCGACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACA 301
QY 1143 TCAGCTGATGACGAGGCTGTCTTTTGGATGATGCTAG 1183
Db 302 TCAGCTGATGACGAGGCTGTCTTTTGGATGATGCTAG 342

CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AA94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 211 BP; 50 A; 51 C; 72 G; 38 T; 0 U; 0 Other;

Query Match 11.9%; Score 141; DB 5; Length 211;
 Best Local Similarity 99.5%; Pred. No. 3.5e-56;
 Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 588 AGATGGAGACTGGTGGACGGTCTGTCTGAAGTCTCAGGCGAGAGTATACATCCCA 647
 DB 20 AGATGGAGACTGGTGGACGGTCTGTCTGAAGTCTCAGGCGAGAGTATACATCCCA 79
 QY 648 GCGTCCACCTGGCGAAAGTCTCCCATGGTGGTCTGTATGAGGGCTGAGCGAGGAGAAAG 707
 DB 80 GCGTCCACCTGGCGAAAGTCTCCCATGGTGGTCTGTATGAGGGCTGAGCGAGGAGAAAG 139
 QY 708 CAGAGAACTGCTGTGTGTACCTGGGAACCTGGAGGGGCTTCTCATCCGGGAGGCC 767
 DB 140 CAGAGAACTGCTGTGTGTACCTGGGAACCTGGAGGGGCTTCTCATCCGGGAGGCC 199
 QY 768 AGACCGAGGAG 779
 DB 200 AGACCGAGGAG 211

RESULT 15
 AA112879
 ID AA112879 standard; DNA; 432 BP.
 AC AA112879;
 XX
 XX
 DT 12-OCT-2001 (first entry)
 DE Probe #2812 for gene expression analysis in human cervical cell sample.
 DE
 DE
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000670.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.
 XX
 PS Claim 25; SEQ ID NO 2812; 487pp; English.
 XX

CC The present invention relates to human single exon nucleic acid probes
 CC (SNP). The present sequence is one such probe. The SNPs are derived
 CC from human HeLa cells. The SNPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging of
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 432 BP; 109 A; 115 C; 111 G; 97 T; 0 U; 0 Other;

Query Match 11.3%; Score 134; DB 4; Length 432;
 Best Local Similarity 100.0%; Pred. No. 6.8e-53;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 929 GAGCTGGCGGATGACATCTGCTCCTACTCAAGGACCCCTGTCTCTGAGAGGGCTGCG 988
 DB 270 GAGCTGGCGGATGACATCTGCTCCTACTCAAGGACCCCTGTCTCTGAGAGGGCTGCG 329
 QY 989 CCGCTCCCTGCCAAGCATATACCCCTACTGTGACCTGTGACGAGGACACCACTCAACTGG 1048
 DB 330 CCGCTCCCTGCCAAGCATATACCCCTACTGTGACCTGTGACGAGGACACCACTCAACTGG 389
 QY 1049 AAAGAGCTGGACAG 1062
 DB 390 AAAGAGCTGGACAG 403

RESULT 16
 ABA54580
 ID ABA54580 standard; DNA; 432 BP.
 AC ABA54580;
 XX
 DT 01-FEB-2002 (first entry)
 DE Human foetal liver single exon nucleic acid probe #2885.
 DE
 DE
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 KW Homo sapiens.
 OS
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000669.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human foetal liver.
 XX
 PS Claim 1; SEQ ID NO 2885; 639pp + Sequence Listing; English.
 XX
 PT The invention relates to a single exon nucleic acid probe for measuring
 PT human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human foetal liver. The

QY 929 GAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTGCAGAGGCTGGC 988
DB 270 GAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTGCAGAGGCTGGC 329
QY 989 CCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGAGCACCCTCAACTGG 1048
DB 330 CCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGAGCACCCTCAACTGG 389
QY 1049 AAAGAGCTGGACAG 1062
DB 390 AAAGAGCTGGACAG 403
RESULT 21
ID AAK02872 standard; DNA; 432 BP.
XX AC AAK02872;
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe SEQ ID NO: 2863.
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX OS ss.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PS WPI; 2001-483446/52.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PS brains.
XX PS Example 4; SEQ ID NO 2863; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system,
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancers. The present sequence is one of the probes of the
XX CC invention
XX SQ Sequence 432 BP; 109 A; 115 C; 111 G; 97 T; 0 U; 0 Other;
Query Match 11.3%; Score 134; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 6.8e-53;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 929 GAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTGCAGAGGCTGGC 988
DB 270 GAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTGCAGAGGCTGGC 329

QY 989 CCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGAGCACCCTCAACTGG 1048
DB 330 CCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGAGCACCCTCAACTGG 389
QY 1049 AAAGAGCTGGACAG 1062
DB 390 AAAGAGCTGGACAG 403
RESULT 22
ID ABS27912 standard; DNA; 432 BP.
XX AC ABS27912;
XX DT 25-FEB-2003 (first entry)
XX DE Human liver single exon probe, SEQ ID No 2902.
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
XX KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KW coronary heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PS WPI; 2001-488898/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PS gene expression in human adult liver.
XX PS Claim 1; SEQ ID NO 2902; 658pp; English.
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 1310y defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (I) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
XX CC involved in genetic liver diseases such as cirrhosis,
XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX CC associated with coronary heart disease. ABS25011-ABS51005 represent human
XX CC liver single exon nucleic acid probes of the invention. Note: The
XX CC sequence information for this patent does not appear in the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 432 BP; 109 A; 115 C; 111 G; 97 T; 0 U; 0 Other;
Query Match 11.3%; Score 134; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 6.8e-53;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 929 GAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTGCAGAGGCTGGC 988

Db 270 GAGCTGGCGGATGACATCTGCTGCTCTACTCAAGGACCCCTGTGTCTCGAGAGGCTGGC 329

Qy 989 CCGCTCCCTGGCAAGGATATACCCCTACTCTGTGACTGTGCGAGGACACCACTCAACTGG 1048

Db 330 CCGCTCCCTGGCAAGGATATACCCCTACTCTGTGACTGTGCGAGGACACCACTCAACTGG 389

Qy 1049 AAAGAGCTGGACAG 1062

Db 390 AAAGAGCTGGACAG 403

RESULT 23

AA102797

ID AA102797 standard; DNA; 432 BP.

XX AA102797;

XX 09-OCT-2001 (first entry)

DT Probe #2788 used to measure gene expression in human breast sample.

DE Probe; human; breast disease; breast cancer; development disorder; ss;

XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

XX WO200157270-A2.

XX 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US000661.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression in a human breast.

XX Claim 25; SEQ ID NO 2788; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes.

XX The present sequence is one such probe. The probes are useful for

XX measuring human gene expression in a human breast sample, where the probe

XX hybridises at high stringency to a nucleic acid expressed in the human

XX breast. The probes are useful for predicting, diagnosing, grading,

XX staging, monitoring and prognosing diseases of the human breast,

XX particularly those diseases with polygenic aetiology. The diseases

XX include: breast cancer, disorders of development, inflammatory diseases

XX of the breast, fibrocystic changes, proliferative breast disease and non-

XX carcinoma tumours. Note: The sequence data for this patent did not form

XX part of the printed specification, but was obtained in electronic format

XX directly from WIPO at ftp.wipo.int/pub/published_pat_sequences

XX SQ Sequence 432 BP; 109 A; 115 C; 111 G; 97 T; 0 U; 0 Other;

Query Match 11.3%; Score 134; DB 5; Length 432;

Best Local Similarity 100.0%; Pred. No. 6.8e-53;

Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 929 GAGCTGGCGGATGACATCTGCTGCTCTACTCAAGGACCCCTGTGTCTCGAGAGGCTGGC 988

Db 270 GAGCTGGCGGATGACATCTGCTGCTCTACTCAAGGACCCCTGTGTCTCGAGAGGCTGGC 329

Qy 989 CCGCTCCCTGGCAAGGATATACCCCTACTCTGTGACTGTGCGAGGACACCACTCAACTGG 1048

Db 330 CCGCTCCCTGGCAAGGATATACCCCTACTCTGTGACTGTGCGAGGACACCACTCAACTGG 389

Qy 1049 AAAGAGCTGGACAG 1062

Db 390 AAAGAGCTGGACAG 403

RESULT 23

ABS02823

ID ABS02823 standard; DNA; 432 BP.

XX ABS02823;

XX 19-AUG-2002 (first entry)

DT Human genome-derived single exon probe from lung SEQ ID NO 2814.

DE Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX chronic obstructive pulmonary disease; interstitial lung disease;

XX familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX primary ciliary dyskinesia; pulmonary hypertension;

XX hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to

XX measure gene expression in human lung samples.

XX Claim 1; SEQ ID NO 2814; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon

XX nucleic acid probes for measuring gene expression in a sample derived

XX from human lung comprising single exon nucleic acid probes having one of

XX 12614 nucleic acid sequences mentioned in the specification, or their

XX complements or the 12397 open reading frames derived from the 12614

XX probes. Also included are a microarray comprising the novel set of probes

XX; the novel set of probes which hybridise at high stringency to a nucleic

XX acid expressed in the human lung; measuring gene expression in a sample

XX derived from human lung, comprising (a) contacting the array with a

XX collection of detectably labeled nucleic acids derived from human lung

XX mRNA, and (b) measuring the label detectably bound to each probe of the

XX array; identifying exons in a eukaryotic genome, comprising (a)

XX algorithmically predicting at least one exon from genomic sequences of

XX the eukaryote; and (b) detecting specific hybridisation of detectably

XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

XX having a fragment identical to the predicted exon, the probe is included

in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences

[illegible]

RESULT 25
AAII4520
ID AAI14520 standard; DNA; 448 BP.
XX
XX AC AAI14520;
XX AC
XX DT 12-OCT-2001 (first entry)
XX DE
XX DE Probe #4453 for gene expression analysis in human cervical cell sample.
XX DE
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer; ss.
XX OS
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX PD
XX PD 09-AUG-2001.
XX PF
XX PF 30-JAN-2001; 2001WO-US0000670.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX

Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-48901/53.
Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human cervical epithelial cells.
Claim 25; SEQ ID NO 4453; 487pp; English.
The present invention relates to human single exon nucleic acid probes
(SNP). The present sequence is one such probe. The SNPs are derived
from human HeLa cells. The SNPs can be used to produce a single exon
microarray, which can be used for measuring human gene expression in a
sample derived from human cervical epithelial cells. By measuring gene
expression, the probes are therefore useful in grading and/or staging of
diseases of the cervix, notably cervical cancer. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 448 BP; 111 A; 120 C; 113 G; 104 T; 0 U; 0 Other;
Query Match 11.3%; Score 134; DB 4; Length 448;
Best Local Similarity 100.0%; Pred. No. 6.8e-53;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 929 GAGTGGCGATGACATCTGCTCCCTACTCAAGAGCCCTGTGCTCTGCAGAGGGTGGC 988
Db 286 GAGTGGCGATGACATCTGCTCCCTACTCAAGAGCCCTGTGCTCTGCAGAGGGTGGC 345
Qy 989 CCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGAGGACACACTCAACTGG 1048
Db 346 CCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGAGGACACACTCAACTGG 405
Qy 1049 AAAGAGCTGGACAG 1062
Db 406 AAAGAGCTGGACAG 419

RESULT 26
ABS04499
ID ABS04499 standard; DNA; 448 BP.
XX
AC ABS04499;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 4490.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
OS Homo sapiens.
XX
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114193/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 PT
 XX Claim 1; SEQ ID NO 4490; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 448 BP; 111 A; 120 C; 113 G; 104 T; 0 U; 0 Other;
 Query Match 11.3%; Score 134; DB 6; Length 448;
 Best Local Similarity 100.0%; Pred. No. 6.8e-53;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 929 GAGCTGGCGATGACATCTGCTGCTACTCAAGAGCCCTGTGCTCTGAGAGGGGTGGC 988
 DB 286 GAGCTGGCGATGACATCTGCTGCTACTCAAGAGCCCTGTGCTCTGAGAGGGGTGGC 345
 QY 989 CCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCGAGGAGACACCACTCAACTGG 1048
 DB 346 CCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCGAGGAGACACCACTCAACTGG 405
 QY 1049 AAAGAGCTGGACAG 1062
 DB 406 AAAGAGCTGGACAG 419
 RESULT 27
 AAI22119

ID AAI22119 standard; DNA; 96 BP.
 XX
 AC AAI22119;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Probe #12052 for gene expression analysis in human cervical cell sample.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000670.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX WPI; 2001-488901/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.
 XX
 PS Claim 25; SEQ ID NO 12052; 487pp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs). The present sequence is one such probe. The SENPs are derived
 CC from human Hela cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging of
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;
 Query Match 8.1%; Score 96; DB 4; Length 96;
 Best Local Similarity 100.0%; Pred. No. 6.6e-35;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 967 CTGTGTCTGTGAGAGGCTGGCCGCTCCCTGCGAAGATATACCCCTACTGTGACTGT 1026
 DB 1 CTGTGTCTGTGAGAGGCTGGCCGCTCCCTGCGAAGATATACCCCTACTGTGACTGT 60
 QY 1027 GCAGAGACACCACTCAACTGGAAGAGCTGGACAG 1062
 DB 61 GCAGAGACACCACTCAACTGGAAGAGCTGGACAG 96
 RESULT 28
 ABA67198
 ID ABA67198 standard; DNA; 96 BP.
 XX
 AC ABA67198;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #15503.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KW Homo sapiens.
OS WO200157277-A2.
PN 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US000669.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
XX Claim 4; SEQ ID NO 15503; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human fetal liver. The
XX present sequence is a single exon nucleic acid probe of the invention.
XX Note: the sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;
SQ

Query Match 8.1%; Score 96; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 967 CTGTGTCCTGCAGAGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGT 1026
DB 1 CTGTGTCCTGCAGAGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGT 60
QY 1027 GCAGAGGACCACTCACTGGAAGAGCTGGACAG 1062
DB 61 GCAGAGGACCACTCACTGGAAGAGCTGGACAG 96

RESULT 29
AAI47414
ID AAI47414 standard; DNA; 96 BP.
XX AAI47414;
AC AAI47414;
DT 17-OCT-2001 (first entry)
XX Probe #16100 used to measure gene expression in human placenta sample.
DE XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX Homo sapiens.
XX WO200157272-A2.
XX 09-AUG-2001.
XX

Query Match 8.1%; Score 96; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 967 CTGTGTCCTGCAGAGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGT 1026
DB 1 CTGTGTCCTGCAGAGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGT 60
QY 1027 GCAGAGGACCACTCACTGGAAGAGCTGGACAG 1062
DB 61 GCAGAGGACCACTCACTGGAAGAGCTGGACAG 96

RESULT 30
ABA49284
ID ABA49284 standard; DNA; 96 BP.
XX ABA49284;
AC ABA49284;
DT 01-FEB-2002 (first entry)
XX Human breast cell single exon nucleic acid probe #7979.
DE XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer; ss.
XX Homo sapiens.
XX WO200157271-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000662.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
XX Claim 4; SEQ ID NO 7979; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;

Query Match 8.1%; Score 96; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 967 CTGTGTCCTGCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGT 1026
Db 1 CTGTGTCCTGCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGT 60

QY 1027 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 1062
Db 61 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 96

RESULT 31
ABA34292
ID ABA34292 standard; DNA; 96 BP.
XX
XX ABA34292;
XX
XX 23-JAN-2002 (first entry)
XX
XX Probe #12758 for gene expression analysis in human heart cell sample.
DE
DE Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.

XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
XX Claim 4; SEQ ID NO 12758; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;
SQ

Query Match 8.1%; Score 96; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 967 CTGTGTCCTGCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGT 1026
Db 1 CTGTGTCCTGCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGT 60

QY 1027 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 1062
Db 61 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 96

RESULT 32
AAK41374
ID AAK41374 standard; DNA; 96 BP.
XX
XX AAK41374;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 15931.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 15931; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;
SQ
Query Match 8.1%; Score 96; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 967 CTGTGTCCTGCAGAGGCTGGCCGCTCCCTGCGAAGGATATACCCCTACCTGTGACTGT 1026
Db 1 CTGTGTCCTGCAGAGGCTGGCCGCTCCCTGCGAAGGATATACCCCTACCTGTGACTGT 60
QY 1027 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 1062
Db 61 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 96
RESULT 33
AAK15640
ID AAK15640 standard; DNA; 96 BP.
XX
XX AAK15640;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 15631.
DE
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000667.
PP
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
PP
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-483446/52.
DR
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
PT
XX Example 4; SEQ ID NO 15631; 650pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;
SQ
Query Match 8.1%; Score 96; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 967 CTGTGTCCTGCAGAGGCTGGCCGCTCCCTGCGAAGGATATACCCCTACCTGTGACTGT 1026
Db 1 CTGTGTCCTGCAGAGGCTGGCCGCTCCCTGCGAAGGATATACCCCTACCTGTGACTGT 60
QY 1027 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 1062
Db 61 GCAGAGGACCACTCAACTGGAAGAGCTGGACAG 96
RESULT 34
ABS40966
ID ABS40966 standard; DNA; 96 BP.
XX
XX ABS40966;
XX
XX 25-FEB-2003 (first entry)
DT
XX Human liver single exon probe, SEQ ID NO 15956.
DE
XX Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000664.
PP
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
PP
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488898/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
PT
XX Claim 4; SEQ ID NO 15956; 658pp; English.
PS
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;
 SQ Query Match 8.1%; Score 96; DB 4; Length 96;
 Best Local Similarity 100.0%; Pred. No. 6.6e-35;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 967 CTGTGTCCTGCAGAGGCTGGCCCGCTCCCTGCGCAAGGATATACCCCTACCTGTGACTGT 1026
 Db 1 CTGTGTCCTGCAGAGGCTGGCCCGCTCCCTGCGCAAGGATATACCCCTACCTGTGACTGT 60
 QY 1027 GCAGAGGACACCACTCAACTGGAAGAGCTGGACAG 1062
 Db 61 GCAGAGGACACCACTCAACTGGAAGAGCTGGACAG 96
 RESULT 35
 ID AAI07818
 AC AAI07818;
 XX 09-OCT-2001 (first entry)
 DE Probe #7809 used to measure gene expression in human breast sample.
 XX Probe; human; breast disease; breast cancer; development disorder; ss;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX Homo sapiens.
 OS
 XX WO200157270-A2.
 XX 09-AUG-2001.
 XX 29-JAN-2001; 2001WO-US000661.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-476286/51.
 DR Novel single exon nucleic acid probe used to measuring gene expression in
 PT a human breast.
 XX Claim 25; SEQ ID NO 7809; 322pp; English.
 XX The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridises at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and non-
 CC carcinoma tumours. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;
 SQ Query Match 8.1%; Score 96; DB 5; Length 96;

Best Local Similarity 100.0%; Pred. No. 6.6e-35;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 967 CTGTGTCCTGCAGAGGCTGGCCCGCTCCCTGCGCAAGGATATACCCCTACCTGTGACTGT 1026
 Db 1 CTGTGTCCTGCAGAGGCTGGCCCGCTCCCTGCGCAAGGATATACCCCTACCTGTGACTGT 60
 QY 1027 GCAGAGGACACCACTCAACTGGAAGAGCTGGACAG 1062
 Db 61 GCAGAGGACACCACTCAACTGGAAGAGCTGGACAG 96
 RESULT 36
 ID ABS15380
 XX ABS15380 standard; DNA; 96 BP.
 AC ABS15380;
 XX 19-AUG-2002 (first entry)
 DE Human genome-derived single exon probe ORF from lung SEQ ID NO 15371.
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.
 XX Homo sapiens.
 OS
 XX WO200186003-A2.
 XX 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US000665.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 DR Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX Claim 4; SEQ ID NO 15371; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC ; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably

labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a single exon probe open reading frame of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 96 BP; 22 A; 29 C; 28 G; 17 T; 0 U; 0 Other;

Query Match 8.1%; Score 96; DB 6; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 967 CTGTCCTGAGAGGCTGGCCGCTCCCTGGCAGGATATACCCCTACCTGTGACTGT 1025
DB 1 CTGTCCTGAGAGGCTGGCCGCTCCCTGGCAGGATATACCCCTACCTGTGACTGT 60

QY 1027 GCAGAGGACCACTCACTGGAAGAGCTGGACAG 1062
DB 61 GCAGAGGACCACTCACTGGAAGAGCTGGACAG 96

RESULT 37
ID AAL44087
XX AAL44087 standard; cDNA; 1348 BP.
AC AAL44087;
XX
DT 03-OCT-2002 (first entry)
XX
DE Mouse modulator of antigen receptor signalling protein coding sequence.
KW Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
KW MARS; tumour suppressor gene; Src-like adaptor protein; SLAP;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 282..1061
FT /tag= a
FT /product= "Mouse MARS protein"
XX
PN WO200242452-A2.
XX
PD 30-MAY-2002.
XX
XX 26-NOV-2001; 2001WO-CA001662.
XX
XX 27-NOV-2000; 2000CA-02324663.
XX
XX (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI McGlade JC, Loreto MP;

XX WPI; 2002-565654/60.
DR P-PSDB; AAO15456.
XX
PT New isolated modulator of antigen receptor signaling protein or its fragment, useful for treating malignant disorders such as myeloid malignancies, autoimmune disorders and myeloproliferative disorders.
PT
PS Claim 10; Fig 1A; 110pp; English.
XX
CC The invention comprises the amino acid and coding sequences of modulator of antigen receptor signalling (MARS) proteins. The MARS protein is a putative tumour suppressor gene and exhibits structural and sequence similarity to the Src-like adaptor protein (SLAP). The MARS DNA and protein sequences of the invention are useful for the treatment of myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune disorders; immunosuppression, myeloproliferative disorders and malignancies related to the de-regulation of tyrosine kinases (e.g. breast cancer). The present cDNA sequence encodes a mouse MARS protein

XX Sequence 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;

Query Match 3.6%; Score 43; DB 6; Length 1348;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GCAGGGCCCTTCCTCATCCGGAGAGCCAGACCCAGGAGAGGCT 782
DB 621 GCAGGGCCCTTCCTCATCCGGAGAGCCAGACCCAGGAGAGGCT 663

RESULT 38
ID AAD43983
XX AAD43983 standard; DNA; 30 BP.
AC AAD43983;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human leukocyte cDNA library screening PCR primer, PY750.
XX
KW Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2; SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer; neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis; psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy; Crohn's disease; systemic lupus erythematosus; tissue/organ rejection; multiple sclerosis; asthma; acute respiratory distress syndrome; pulmonary disorder; dermatological; neuroprotective; PCR; primer; ss.
XX
OS Homo sapiens.
XX
XX WO200242457-A1.
XX
PD 30-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-US043367.
XX
PR 22-NOV-2000; 2000US-0252545P.
XX
FA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;
XX
XX WPI; 2002-463632/49.
XX
XX Novel substantially purified human SH2/SH3-domain-containing adapter polypeptide, termed Src-like Adapter Protein-2, useful for therapeutic intervention in immunological and inflammatory disorders and cancer.
XX
XX Example 1; Page 60; 85pp; English.
XX
CC The invention relates to a substantially purified human SH2/SH3-domain-containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-

2). The invention is useful for treating an immune disorder involving hyperactivity of B- or T- lymphocytes in a mammal. The invention is useful for screening for antagonists or inhibitors of the interaction of hSLAP-2 with cellular signalling compounds, for diagnosing, treating or preventing diseases or disorders associated with aberrant or uncontrolled cellular signal transduction, for determining those cellular signalling molecules which associate with hSLAP-2 and which provide critical signals for cell activation, and as effectors in methods to affect T-cell activation. The invention is useful in screening assays to identify and detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for potential use to treat autoimmune diseases which may be caused by hyperactivated B cells, as well as to treat diseases which may be caused by hyperactivated T cells, in addition to other immune system related conditions, diseases, or disorders, T-cell and B-cell neoplasms, inflammation disorders, diseases and conditions, rheumatoid arthritis, osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's and ulcerative colitis), allergies, particularly those involving hyperactivity of B-cells and T- cells, or other immune cells, such as mast cells or eosinophils, autoimmune diseases such as systemic lupus erythematosus and multiple sclerosis, pulmonary diseases including asthma, acute respiratory distress syndrome, and chronic obstructive pulmonary disorder, tissue/ organ rejection and cancer. The invention is useful in gene therapy. The present sequence is human leukocyte cDNA library screening PCR primer

Sequence 30 BP; 5 A; 8 C; 9 G; 8 T; 0 U; 0 Other;

Query Match 2.5%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

923 TACTCTGAGCTGGCGATGACATCTCTGTC 952
1 TACTCTGAGCTGGCGATGACATCTCTGTC 30

RESULT 39
ID ABK61506/c
XX ABK61506 standard; DNA; 26 BP.

XX ABK61506;
XX
XX 19-JUN-2002 (first entry)
XX Human NOV13 RT-PCR probe.

XX Human; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
XX cell signal processing disorder; metabolic pathway modulation disorder;
XX diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; prober;
XX uterus cancer; immune response; graft-versus-host disease; Exon linking;
XX acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
XX hypertension; congenital heart defects; multiple sclerosis; inflammation;
XX Albright hereditary osteodystrophy; reverse transcriptase PCR.

XX Homo sapiens.
XX WO200216599-A2.

XX 28-FEB-2002.

XX 27-AUG-2001; 2001WO-US026510.

XX 25-AUG-2000; 2000US-0228191P.

XX 08-FEB-2001; 2001US-0267300P.

XX 20-FEB-2001; 2001US-0269961P.

XX 20-MAR-2001; 2001US-0277337P.

XX (CURA-) CURAGEN CORP.

XX (CORT-) COR THERAPEUTICS INC.

XX Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;
XX Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;

DR WPI; 2002-280937/32.
XX
XX New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
XX Example 2; Page 234; 263pp; English.

XX The invention relates to an isolated polypeptide (NOVX) a mature form of
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide
CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it
CC and antibody against it, are useful for treating or preventing (e.g. by
CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC atherosclerosis, a disorder related to cell signal processing and
CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC and nucleic acids are also useful for determining the presence of
CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are
CC especially useful in therapeutic or prophylactic applications for
CC disorders associated with aberrant NOVX expression or activity, e.g.
CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC cancer), immune response, graft-versus-host disease, acquired
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC congenital heart defects, multiple sclerosis, inflammation or Albright
CC hereditary osteodystrophy and many other diseases listed in the
CC specification. The DNA encoding the protein is useful in gene therapy for
CC treating the conditions. This is also useful in detection assays,
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC human disorders, as well as in diagnostic applications. The present
CC sequence is a reverse transcriptase (RT)-PCR probe used to measure tissue
CC specific expression of mRNA encoding a NOVX protein

XX Sequence 26 BP; 3 A; 8 C; 6 G; 9 T; 0 U; 0 Other;

Query Match 2.2%; Score 26; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 AAGGACACTGGCAGACTTCCAGAAGG 276
Db 26 AAGGACACTGGCAGACTTCCAGAAGG 1

RESULT 40
AAL44100/c
ID AAL44100 standard; DNA; 25 BP.

XX AAL44100;

XX 03-OCT-2002 (first entry)

XX Human modulator of antigen receptor signalling protein PCR primer 4.
XX Human; ss; gene therapy; modulator of antigen receptor signalling; MARS;
XX tumour suppressor gene; Src-like adaptor protein; SLAP; PCR;
XX myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
XX immunosuppression; myeloproliferative disorder; breast cancer; primer.

XX Homo sapiens.

XX WO200242452-A2.

XX 30-MAY-2002.

XX 26-NOV-2001; 2001WO-CA001662.

XX 27-NOV-2000; 2000CA-02324663.

XX (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX Mcglade JC, Loreto MP;

XX WPI; 2002-566564/60.

```
XX New isolated modulator of antigen receptor signaling protein or its
PT fragment, useful for treating malignant disorders such as myeloid
PT malignancies, autoimmune disorders and myeloproliferative disorders.
PS
XX Claim 20; Page 69; 110pp; English.
XX
CC The invention comprises the amino acid and coding sequences of modulator
CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC putative tumour suppressor gene and exhibits structural and sequence
CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
CC protein sequences of the invention are useful for the treatment of
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC disorders, immunosuppression, myeloproliferative disorders and
CC malignancies related to the de-regulation of tyrosine kinases (e.g.
CC breast cancer). The present DNA sequence represents a human MARS protein
CC PCR primer
XX
SQ Sequence 25 BP; 9 A; 8 C; 5 G; 3 T; 0 U; 0 Other;

Query Match      2.1%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1159 GGCTGCTCTTTGGATGATGCCTAG 1183
Db 25 GGCTGCTCTTTGGATGATGCCTAG 1

RESULT 41
AAL44098/c
ID AAL44098 standard; DNA; 25 BP.
XX
AC AAL44098;
XX
XX 03-OCT-2002 (first entry)
XX
DE Human modulator of antigen receptor signalling protein PCR primer 2.
XX
KW Human; ss; gene therapy; modulator of antigen receptor signalling; MARS;
KW tumour suppressor gene; Scr-like adaptor protein; SLAP; PCR;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer; primer.
XX
OS Homo sapiens.
XX
XX WO200242452-A2.
XX
XX 30-MAY-2002.
XX
XX 26-NOV-2001; 2001WO-CA001662.
XX
XX 27-NOV-2000; 2000CA-02324663.
XX
XX (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
XX Mcglade JC, Loreto MP;
XX
XX WPI; 2002-566564/60.
XX
XX New isolated modulator of antigen receptor signaling protein or its
XX fragment, useful for treating malignant disorders such as myeloid
XX malignancies, autoimmune disorders and myeloproliferative disorders.
XX
XX Claim 20; Page 69; 110pp; English.
XX
XX The invention comprises the amino acid and coding sequences of modulator
XX of antigen receptor signalling (MARS) proteins. The MARS protein is a
XX putative tumour suppressor gene and exhibits structural and sequence
XX similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
XX protein sequences of the invention are useful for the treatment of
XX myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
XX disorders, immunosuppression, myeloproliferative disorders and
XX malignancies related to the de-regulation of tyrosine kinases (e.g.
XX breast cancer). The present DNA sequence represents a human MARS protein
XX PCR primer
XX
SQ Sequence 25 BP; 9 A; 8 C; 5 G; 3 T; 0 U; 0 Other;

Query Match      2.1%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1159 GGCTGCTCTTTGGATGATGCCTAG 1183
Db 25 GGCTGCTCTTTGGATGATGCCTAG 1

RESULT 42
AAL44099
ID AAL44099 standard; DNA; 25 BP.
XX
AC AAL44099;
XX
XX 03-OCT-2002 (first entry)
XX
DE Human modulator of antigen receptor signalling protein PCR primer 3.
XX
KW Human; ss; gene therapy; modulator of antigen receptor signalling; MARS;
KW tumour suppressor gene; Scr-like adaptor protein; SLAP; PCR;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer; primer.
XX
OS Homo sapiens.
XX
XX WO200242452-A2.
XX
XX 30-MAY-2002.
XX
XX 26-NOV-2001; 2001WO-CA001662.
XX
XX 27-NOV-2000; 2000CA-02324663.
XX
XX (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
XX Mcglade JC, Loreto MP;
XX
XX WPI; 2002-566564/60.
XX
XX New isolated modulator of antigen receptor signaling protein or its
XX fragment, useful for treating malignant disorders such as myeloid
XX malignancies, autoimmune disorders and myeloproliferative disorders.
XX
XX Claim 20; Page 69; 110pp; English.
XX
XX The invention comprises the amino acid and coding sequences of modulator
XX of antigen receptor signalling (MARS) proteins. The MARS protein is a
XX putative tumour suppressor gene and exhibits structural and sequence
XX similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
XX protein sequences of the invention are useful for the treatment of
XX myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
XX disorders, immunosuppression, myeloproliferative disorders and
XX malignancies related to the de-regulation of tyrosine kinases (e.g.
XX breast cancer). The present DNA sequence represents a human MARS protein
XX PCR primer
XX
SQ Sequence 25 BP; 9 A; 8 C; 5 G; 3 T; 0 U; 0 Other;

Query Match      2.1%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 773 AGGAGAGGCTTCTTACTCTCTGTCAG 797
Db 1 AGGAGAGGCTTCTTACTCTCTGTCAG 25
```

```

RESULT 43
AAL44097
ID AAL44097 standard; DNA; 23 BP.
XX
AC AAL44097;
XX
DT 03-OCT-2002 (first entry)
XX
DE Human modulator of antigen receptor signalling protein PCR primer 1.
XX
KW Human; ss; gene therapy; modulator of antigen receptor signalling; MARS;
KW tumour suppressor gene; scr-like adaptor protein; SLAP; PCR;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer; primer.
XX
OS Homo sapiens.
XX
FN WO200242452-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-CA001662.
XX
PR 27-NOV-2000; 2000CA-02324663.
XX
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Mcglade JC, Loreto MP;
XX
PS WPI; 2002-566564/60.
XX
CC New isolated modulator of antigen receptor signaling protein or its
CC fragment, useful for treating malignant disorders such as myeloid
CC malignancies, autoimmune disorders and myeloproliferative disorders.
XX
PS Claim 20; Page 67; 110pp; English.
XX
CC The invention comprises the amino acid and coding sequences of modulator
CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC putative tumour suppressor gene and exhibits structural and sequence
CC similarity to the scr-like adaptor protein (SLAP). The MARS DNA and
CC protein sequences of the invention are useful for the treatment of
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC disorders, immunosuppression, myeloproliferative disorders and
CC malignancies related to the de-regulation of tyrosine kinases (e.g.
CC breast cancer). The present DNA sequence represents a human MARS protein
CC PCR primer
XX
SQ Sequence 23 BP; 7 A; 5 C; 8 G; 3 T; 0 U; 0 Other;
Query Match 1.9%; Score 23; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 398 ATGGGAAGTCTGCCAGCAGAG 420
Db 1 ATGGGAAGTCTGCCAGCAGAG 23
|||||
RESULT 44
ABK61507
ID ABK61507 standard; DNA; 22 BP.
XX
AC ABK61507;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human NOV13 RT-PCR primer #2.
XX
KW Human; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
KW cell signal processing disorder; metabolic pathway modulation disorder;
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; primer;
XX

```

```

KW uterus cancer; immune response; graft-versus-host disease; Exon linking;
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;
KW Albright hereditary osteodystrophy; reverse transcriptase PCR.
XX
OS Homo sapiens.
XX
FN WO200216599-A2.
XX
PD 28-FEB-2002.
XX
PR 27-AUG-2001; 2001WO-US026510.
XX
PR 25-AUG-2000; 2000US-0228191P.
XX
PR 08-FEB-2001; 2001US-0267300P.
XX
PR 20-FEB-2001; 2001US-0269961P.
XX
PR 20-MAR-2001; 2001US-0277337P.
XX
PA (CURA-) CURAGEN CORP.
XX
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;
XX
DR WPI; 2002-280937/32.
XX
CC New polypeptides for treating or preventing a disorder associated with
CC them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
XX Example 2; Page 234; 263pp; English.
XX
CC The invention relates to an isolated polypeptide (NOVX) a mature form of
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide
CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it,
CC and antibody against it, are useful for treating or preventing (e.g. by
CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC atherosclerosis, a disorder related to cell signal processing and
CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC and nucleic acids are also useful for determining the presence of
CC predispotion to the diseases. The NOVX nucleic acid and polypeptide are
CC especially useful in therapeutic or prophylactic applications for
CC disorders associated with aberrant NOVX expression or activity, e.g.
CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC cancer), immune response, graft-versus-host disease, acquired
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC congenital heart defects, multiple sclerosis, inflammation or Albright
CC hereditary osteodystrophy and many other diseases listed in the
CC specification. The DNA encoding the protein is useful in gene therapy for
CC treating the conditions. This is also useful in detection assays,
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC for developing a powerful assay system for functional analysis of various
CC human disorders, as well as in diagnostic applications. The present
CC sequence is a reverse transcriptase (RT)-PCR primer used to measure
CC tissue specific expression of mRNA encoding a NOVX protein
XX
SQ Sequence 22 BP; 4 A; 3 C; 8 G; 7 T; 0 U; 0 Other;
Query Match 1.9%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 224 TGAGAGAGTTCTGGGTCTCTTA 245
Db 1 TGAGAGAGTTCTGGGTCTCTTA 22
|||||
RESULT 45
AAK55277/c
ID AAK55277 standard; cDNA to mRNA; 3070 BP.
XX
AC AAK55277;
XX

```

DT 08-JUL-1999 (first entry)
 XX Nucleotide sequence of human pink eye dilute gene.
 DE Melanocyte; pink eye dilute gene; melanin; human; ss.
 KW Homo sapiens.
 OS JP11103864-A.
 PN 20-APR-1999.
 PD 02-OCT-1997; 97JP-00286143.
 PF 02-OCT-1997; 97JP-00286143.
 PR (POKK) POLA CHEM IND INC.
 PA WPI; 1999-305841/26.
 DR Discrimination of melanocyte - using a pink eye dilute gene-related
 PT substance as the index.
 PT Claim 2; Page 3-4; 5pp; Japanese.
 PS The invention relates to the discrimination of melanocyte using a pink
 CC eye dilute gene-related substance as the index. A method is also provided
 CC for the evaluation of a melanin production inhibitor in which the
 CC activity of melanocyte is discriminated by the above method and the
 CC reduction in activity is used as the index. The pink eye dilute gene-
 CC related substance is a new factor participating to melanin production.
 CC The present sequence represents the nucleotide sequence of human pink eye
 CC dilute gene
 XX Sequence 3070 BP; 694 A; 790 C; 808 G; 778 T; 0 U; 0 Other;
 SQ Query Match 1.9%; Score 22; DB 2; Length 3070;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 316 TCAGCAGAGCTGCTCTCCCAAG 337
 DB 450 TCAGCAGAGCTGCTCTCCCAAG 429
 RESULT 47
 AAD43981
 ID AAD43981 standard; DNA; 21 BP.
 AC AAD43981;
 XX 13-DEC-2002 (first entry)
 DT Human leukocyte cDNA library screening PCR primer, PY749.
 DE Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;
 KW pulmonary disorder; dermatological; neuroprotective; PCR; primer; ss.
 XX Homo sapiens.
 OS WO200242457-A1.
 PN 30-MAY-2002.
 PD 20-NOV-2001; 2001WO-US043367.
 PF 22-NOV-2000; 2000US-0252545P.
 PR (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;
 PI WPI; 2002-463632/49.
 XX Novel substantially purified human SH2/SH3-domain-containing adapter
 XX polypeptide, termed Src-like Adapter Protein-2, useful for therapeutic
 XX intervention in immunological and inflammatory disorders and cancer.
 PS Example 1; Page 60; 85pp; English.
 XX The invention relates to a substantially purified human SH2/SH3-domain-
 CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-
 CC 2). The invention is useful for treating an immune disorder involving
 CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is
 CC useful for screening for antagonists or inhibitors of the interaction of
 CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or
 CC preventing diseases or disorders associated with aberrant or uncontrolled
 CC cellular signal transduction, for determining those cellular signalling
 CC molecules which associate with hSLAP-2 and which provide critical signals
 CC for cell activation, and as effectors in methods to affect T- cell

DT 08-JUL-1999 (first entry)
 XX Nucleotide sequence of human pink eye dilute gene.
 DE Melanocyte; pink eye dilute gene; melanin; human; ss.
 KW Homo sapiens.
 OS JP11103864-A.
 PN 20-APR-1999.
 PD 02-OCT-1997; 97JP-00286143.
 PF 02-OCT-1997; 97JP-00286143.
 PR (POKK) POLA CHEM IND INC.
 PA WPI; 1999-305841/26.
 DR Discrimination of melanocyte - using a pink eye dilute gene-related
 PT substance as the index.
 PT Claim 2; Page 3-4; 5pp; Japanese.
 PS The invention relates to the discrimination of melanocyte using a pink
 CC eye dilute gene-related substance as the index. A method is also provided
 CC for the evaluation of a melanin production inhibitor in which the
 CC activity of melanocyte is discriminated by the above method and the
 CC reduction in activity is used as the index. The pink eye dilute gene-
 CC related substance is a new factor participating to melanin production.
 CC The present sequence represents the nucleotide sequence of human pink eye
 CC dilute gene
 XX Sequence 3070 BP; 694 A; 790 C; 808 G; 778 T; 0 U; 0 Other;
 SQ Query Match 1.9%; Score 22; DB 2; Length 3070;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 316 TCAGCAGAGCTGCTCTCCCAAG 337
 DB 450 TCAGCAGAGCTGCTCTCCCAAG 429
 RESULT 46
 AAX55278/c
 ID AAX55278 standard; mRNA; 3070 BP.
 AC AAX55278;
 XX 08-JUL-1999 (first entry)
 DT mRNA sequence of human pink eye dilute gene.
 DE Melanocyte; pink eye dilute gene; melanin; human; ss.
 KW Homo sapiens.
 OS JP11103864-A.
 PN 20-APR-1999.
 PD 02-OCT-1997; 97JP-00286143.
 PF 02-OCT-1997; 97JP-00286143.
 PR (POKK) POLA CHEM IND INC.
 PA WPI; 1999-305841/26.
 DR Discrimination of melanocyte - using a pink eye dilute gene-related
 PT substance as the index.
 PT

CC activation. The invention is useful in screening assays to identify and
 CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for
 CC potential use to treat autoimmune diseases which may be caused by
 CC hyperactivated B cells, as well as to treat diseases which may be caused
 CC by hyperactivated T cells, in addition to other immune system related
 CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,
 CC inflammation disorders, diseases and conditions, rheumatoid arthritis,
 CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's
 CC and ulcerative colitis), allergies, particularly those involving
 CC hyperactivity of B-cells and T-cells, or other immune cells, such as
 CC mast cells or eosinophils, autoimmune diseases such as systemic lupus
 CC erythematosus and multiple sclerosis, pulmonary diseases including
 CC asthma, acute respiratory distress syndrome, and chronic obstructive
 CC pulmonary disorder, tissue/organ rejection and cancer. The invention is
 CC useful in gene therapy. The present sequence is human leukocyte cDNA
 CC library screening PCR primer

XX
 SQ Sequence 21 BP; 7 A; 6 C; 5 G; 3 T; 0 U; 0 Other;

Query Match 1.8%; Score 21; DB 6; Length 21;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 827 CGGATCAGACACTACAGGATC 847

Db 1 CGGATCAGACACTACAGGATC 21

RESULT 48

ID ABK61505/c
 AC ABK61505; DNA; 20 BP.

AC ABK61505;

DT 18-JUN-2002 (first entry)

DE Human NOV13 RT-PCR primer #1.

XX Human; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
 KW cell signal processing disorder; metabolic pathway modulation disorder;
 KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; primer;
 KW uterus cancer; immune response; graft-versus-host disease; Exon linking;
 KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
 KW hypertension; congenital heart defects; multiple sclerosis; inflammation;
 KW Albright hereditary osteodystrophy; reverse transcriptase PCR.

XX Homo sapiens.

XX WC200216599-A2.

XX 28-FEB-2002.

XX 27-AUG-2001; 2001WO-US026510.

XX 25-AUG-2000; 2000US-0228191P.

PR 08-FEB-2001; 2001US-0267300P.

PR 20-FEB-2001; 2001US-0269961P.

PR 20-MAR-2001; 2001US-0277337P.

XX (CURA-) CURAGEN CORP.

PA (CORT-) COR THERAPEUTICS INC.

XX Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;

PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;

XX WPI; 2002-280937/32.

XX New polypeptides for treating or preventing a disorder associated with

PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.

XX Example 2; Page 234; 263pp; English.

XX The invention relates to an isolated polypeptide (NOVX) a mature form of

CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide
 CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
 CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it
 CC and antibody against it, are useful for treating or preventing (e.g. by
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
 CC atherosclerosis, a disorder related to cell signal processing and
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
 CC and nucleic acids are also useful for determining the presence of
 CC predilection to the diseases. The NOVX nucleic acid and polypeptide are
 CC especially useful in therapeutic or prophylactic applications for
 CC disorders associated with aberrant NOVX expression or activity, e.g.
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
 CC cancer), immune response, graft-versus-host disease, acquired
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
 CC congenital heart defects, multiple sclerosis, inflammation or Albrit
 CC hereditary osteodystrophy and many other diseases listed in the
 CC specification. The DNA encoding the protein is useful in gene therapy for
 CC treating the conditions. This is also useful in detection assays,
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or
 CC for developing a powerful assay system for functional analysis of various
 CC human disorders, as well as in diagnostic applications. The present
 CC sequence is a reverse transcriptase (RT)-PCR primer used to measure
 CC tissue specific expression of mRNA encoding a NOVX protein

XX Sequence 20 BP; 3 A; 3 C; 8 G; 6 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 CAAAGCCCTAACCTGTCCAG 301

Db 20 CAAAGCCCTAACCTGTCCAG 1

RESULT 49

AAD43982/c

ID AAD43982 standard; DNA; 20 BP.

AC AAD43982;

DT 13-DEC-2002 (first entry)

XX Human leukocyte cDNA library screening PCR primer, PY751.

XX Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;
 KW neoplasia; inflammation disorder; rheumatoid arthritis; osteoarthritis;
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;
 KW pulmonary disorder; dermatological; neuroprotective; PCR; primer; ss.

XX Homo sapiens.

XX WO200242457-A1.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US043367.

XX 22-NOV-2000; 2000US-0252545P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;

XX WPI; 2002-463632/49.

XX Novel substantially purified human SH2/SH3-domain-containing adapter
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic
 PT intervention in immunological and inflammatory disorders and cancer.

PS Example 1; Page 60; 85pp; English.

XX The invention relates to a substantially purified human SH2/SH3-domain-
CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-
CC 2). The invention is useful for treating an immune disorder involving
CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is
CC useful for screening for antagonists or inhibitors of the interaction of
CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or
CC preventing diseases or disorders associated with aberrant or uncontrolled
CC cellular signal transduction, for determining those cellular signalling
CC molecules which associate with hSLAP-2 and which provide critical signals
CC for cell activation, and as effectors in methods to affect T- cell
CC activation. The invention is useful in screening assays to identify and
CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for
CC potential use to treat autoimmune diseases which may be caused by
CC hyperactivated B cells, as well as to treat diseases which may be caused
CC by hyperactivated T cells, in addition to other immune system related
CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,
CC inflammation disorders, diseases and conditions, rheumatoid arthritis,
CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's
CC and ulcerative colitis), allergies, particularly those involving
CC hyperactivity of B-cells and T cells, or other immune cells, such as
CC mast cells or eosinophils, autoimmune diseases such as systemic lupus
CC erythematosus and multiple sclerosis, pulmonary diseases including
CC asthma, acute respiratory distress syndrome, and chronic obstructive
CC pulmonary disorder, tissue/ organ rejection and cancer. The invention is
CC useful in gene therapy. The present sequence is human leukocyte cDNA
CC library screening PCR primer

XX SQ Sequence 20 BP; 4 A; 4 C; 6 G; 6 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1138 CTACATCAGCCGTAATGACG 1157
Db 20 CTACATCAGCCGTAATGACG 1

RESULT 50
AAC21471/c
ID AAC21471 standard; cDNA; 277 BP.
XX AC AAC21471;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 25546.
DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX Gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-00200610.
XX 26-FEB-1999; 99US-0122487P.
XX (GEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.

PS Claim 1; SEQ ID NO 25546; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors

XX SQ Sequence 277 BP; 86 A; 56 C; 75 G; 60 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 3; Length 277;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 210 AGATCCTCCAGGCTGAGAG 229
Db 221 AGATCCTCCAGGCTGAGAG 202

RESULT 51
AAH05869
ID AAH05869 standard; cDNA; 525 BP.
XX AC AAH05869;
XX 26-JUN-2001 (first entry)
XX Human cDNA clone (5'-primer) SEQ ID NO:2704.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX Claim 1; SEQ ID NO 2704; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesising 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB32446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX
 SQ Sequence 525 BP; 149 A; 138 C; 109 G; 116 T; 0 U; 13 Other;

Query Match 1.7%; Score 20; DB 4; Length 525;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 GAAGAAATCTCTGCCAAGC 436
 |||||
 Db 452 GAAGAAATCTCTGCCAAGC 471

RESULT 52

AA163421
 ID AAT63421 standard; DNA; 675 BP.

XX AC AAT63421;

DT 20-JUN-1997 (first entry)

DE FKBP-LCK:SH2 fusion protein DNA.

XX FKBP-LCK:SH2; FK506 binding protein; SH2 domain; Src homology 2;
 XX fusion protein; high throughput assay; ligand; signal transduction;
 KW microscintillation; ss.

XX Homo sapiens.

XX WO9710253-A1.

XX 20-MAR-1997.

XX 11-SEP-1996; 96WO-US014567.

XX 15-SEP-1995; 95US-0003819P.

XX 12-MAR-1996; 96GB-00005210.

XX (MERI) MERCK & CO INC.

XX Marcy A, Salowe SP, Wisniewski D;

XX WPI; 1997-202171/18.

XX P-PSDB; AAW14788.

XX Screening compounds for binding to fusion proteins with defined ligands -
 XX allows high capacity assays and identification of (ant)agonists or
 XX inhibitors for drug development.

XX Claim 29; Page 17-18; 36pp; English.

XX Isolated DNA sequences (AAT63419-21) respectively code for fusion
 CC proteins FKBP-ZAP:SH2, FKBP-SYK:SH2 and FKBP-LCK:SH2 (AAW14786-88)
 CC comprising FK506 binding protein (FKBP) linked via a peptide linker to a
 CC target protein composed of a multiple signal transduction domain, i.e.
 CC ZAP:SH2, SYK:SH2 or LCK:SH2. A DNA fragment encoding the linker is joined
 CC to human FKBP DNA from which the stop codon has been removed. The DNA

CC construct is cloned into an expression vector and target protein DNA is
 CC then added to obtain expression vectors suitable for producing the fusion
 CC proteins in transformed host (pref. E. coli) cells. The fusion proteins
 CC are used in novel high throughput assays utilising microscintillation
 CC plate technology for functional assay of ligand binding

XX SQ Sequence 675 BP; 157 A; 194 C; 191 G; 133 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 2; Length 675;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CTTCTCTATCCGGGAGAC 766
 |||||

Db 443 CTTCTCTATCCGGGAGAC 462

RESULT 53

AA15151
 ID AAX15151 standard; DNA; 675 BP.

XX AC AAX15151;

XX 21-APR-1999 (first entry)

XX DNA encoding a fusion protein of FKBP-Lck.

XX Fusion protein; FK506 binding protein; FKBP; SH2 domain; human Lck;
 XX screening; protein binding; ligand-protein interaction;
 KW protein-protein interaction; protease inhibitor; ss.

XX Synthetic.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..675
 FT /*tag= a

XX WO9841866-A1.

XX 24-SEP-1998.

XX 10-MAR-1998; 98WO-US004610.

XX 14-MAR-1997; 97US-0040795P.

XX (MERI) MERCK & CO INC.

XX Hermes JD, Salowe SP, Sinclair PJ;

XX WPI; 1999-070061/06.

XX P-PSDB; AAW96823.

XX High throughput screening assay - for screening compounds capable of
 XX binding to a fusion protein consisting of, e.g., a target protein and an
 XX FK506-binding protein.

XX Disclosure; Page 23; 42pp; English.

XX The present sequence encodes a fusion protein comprising FK506 binding
 CC protein (FKBP) and the SH2 domain of human Lck. The protein is used to
 CC exemplify the method of the invention. The specification describes a
 CC method for screening for compounds capable of binding to a fusion
 CC protein. The method comprises mixing a test compound, a biotinylated
 CC ligand, the fusion protein, a donor-labelled ligand and acceptor-labelled
 CC streptavidin, incubating the mixture, measuring the time-resolved
 CC fluorescence attributable to the binding of the biotinylated ligand to
 CC the fusion protein in the presence of the test compound and determining
 CC the binding of the biotinylated ligand to the fusion protein in the
 CC presence of the test compound relative to a control assay run in the
 CC absence of the test compound. The methods may be used to determine if
 CC compounds are capable of binding to a protein or are capable of blocking
 CC ligand-protein or protein-protein interactions. They may be used to

CC identify compounds which are protease inhibitors
 XX Sequence 675 BP; 157 A; 194 C; 191 G; 133 T; 0 U; 0 Other;
 SQ

Query Match 1.7%; Score 20; DB 2; Length 675;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCTCATCCGGGAGAGC 766
 |||||
 Db 443 CCTTCTCATCCGGGAGAGC 462

RESULT 54
 ADA69981
 ID ADA69981 standard; DNA; 1047 BP.
 XX
 AC ADA69981;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Rice gene, SEQ ID 3304.
 XX
 KW Plant; bacterial infection; fungal infection; viral infection; rice;
 KW Gene; ds.
 XX
 OS Oryza sativa.
 XX
 PN WO200300898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-IB001105.
 XX
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX
 DR WPI; 2003-175290/17.
 XX
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 PS Claim 6; SEQ ID NO 3304; 899pp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 1047 BP; 163 A; 342 C; 344 G; 195 T; 0 U; 3 Other;

Query Match 1.7%; Score 20; DB 7; Length 1047;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 538 CCGGCGGAGCTGTGCTGA 557
 |||||
 Db 537 CCGGCGGAGCTGTGCTGA 556

RESULT 55

AAQ13983
 ID AAQ13983 standard; DNA; 1254 BP.
 XX
 AC AAQ13983;
 XX
 DT 13-DEC-1991 (first entry)
 XX
 DE Lck gene fused with part of beta-galactosidase gene.
 XX
 KW Multi-cloning site; ss.
 XX
 OS Synthetic.
 XX
 FH Location/Qualifiers
 FT misc_RNA
 FT 1..78
 FT /*tag= a
 FT /note= "beta-galactosidase gene fragment"
 FT
 FT misc_RNA
 FT 79..1254
 FT /*tag= b
 FT /note= "lck gene"
 XX
 PN JF03201994-A.
 XX
 PD 03-SEP-1991.
 XX
 PF 28-DEC-1989; 89JP-00338268.
 XX
 PR 28-DEC-1989; 89JP-00338268.
 XX
 PA (TOKU) TOKUYAMA SODA KK.
 XX
 DR WPI; 1991-300980/41.
 DR P-PSDB; AAR14201.
 XX
 PT Fused polypeptide - has amino acid sequence of beta-galactosidase with a
 PT Lck gene conjugated to the N-terminal via DNA having multi-cloning site.
 XX
 PS Disclosure; Fig 4.2; 15pp; Japanese.
 XX
 CC The sequence consists of the first 78 bp encoding the N-terminal amino
 CC acids of the beta-galactosidase gene fused with the lck gene. It is
 CC prep'd. by a claimed process in which a DNA contg. the lck gene is
 CC inserted into an E.coli expression vector. The vector has DNA contg. part
 CC or all of the beta-galactosidase gene at the appropriate site of the
 CC multi-cloning site. It is useful for producing an antibody specifically
 CC immunoreactive with only a lck gene-derived polypeptide in F cells. The
 CC antibody may recognise lck gene-derived polypeptides in human cells
 XX
 SQ Sequence 1254 BP; 291 A; 361 C; 365 G; 237 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 2; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCTCATCCGGGAGAGC 766
 |||||
 Db 170 CCTTCTCATCCGGGAGAGC 189

RESULT 56
 ADA02968
 ID ADA02968 standard; cDNA; 1530 BP.
 XX
 AC ADA02968;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human LCK carcinoma associated coding sequence, SEQ ID NO:1486.
 XX
 KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW Gene; ss.
 XX

OS Homo sapiens.
 XX WO2003057146-A2.
 XX 17-JUL-2003.
 XX 26-DEC-2002; 2002WO-US041414.
 XX 26-DEC-2001; 2001US-00035832.
 XX (SAGR-) SAGRES DISCOVERY.
 XX Morris DW;
 XX WPI; 2003-587068/55.
 XX New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 XX Claim 1; SEQ ID NO 1486; 245pp; English.
 XX The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed human CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 1530 BP; 352 A; 438 C; 451 G; 289 T; 0 U; 0 Other;
 SQ Query Match 1.7%; Score 20; DB 8; Length 1530;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 747 CCTTCTCATCCGGGAGAC 766
 DB 449 CCTTCTCATCCGGGAGAC 468
 RESULT 57
 ADB72706
 ID ADB72706 standard; cDNA; 1530 BP.
 XX ADB72706;
 XX 04-DEC-2003 (first entry)
 XX Human LCK cDNA.
 XX human; ss; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
 XX cancer; neoplasm; adenocarcinoma; sarcoma; gene.
 XX Homo sapiens.
 XX WO2003008583-A2.
 XX 30-JAN-2003.
 PD

XX 26-DEC-2001; 2001WO-US051291.
 XX 02-MAR-2001; 2001US-00798586.
 XX 23-OCT-2001; 2001US-00004113.
 XX 08-NOV-2001; 2001US-00052482.
 XX 30-NOV-2001; 2001US-00997722.
 XX 20-DEC-2001; 2001US-00034650.
 XX (SAGR-) SAGRES DISCOVERY.
 XX Morris DW, Engelhard EK;
 XX WPI; 2003-239337/23.
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX Claim 1; SEQ ID NO 534; 2304pp; English.
 XX The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined
 CC in the specification. A polynucleotide of the invention has cytostatic
 CC activity, and may have a use in gene therapy, or in a vaccine. The
 CC recombinant nucleic acids and polypeptides are useful for treating
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 CC sarcomas. The present sequence represents a human cDNA of the invention.
 XX
 XX Sequence 1530 BP; 352 A; 438 C; 451 G; 289 T; 0 U; 0 Other;
 SQ Query Match 1.7%; Score 20; DB 9; Length 1530;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 747 CCTTCTCATCCGGGAGAC 766
 DB 449 CCTTCTCATCCGGGAGAC 468
 RESULT 58
 ADC85448
 ID ADC85448 standard; DNA; 1530 BP.
 XX ADC85448;
 XX 01-JAN-2004 (first entry)
 XX Human Lck coding sequence.
 XX Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
 XX secreted; transmembrane; intracellular; ds.
 XX Homo sapiens.
 XX WO2003045230-A2.
 XX 05-JUN-2003.
 XX 02-DEC-2002; 2002WO-US038582.
 XX 30-NOV-2001; 2001US-00997722.
 XX (SAGR-) SAGRES DISCOVERY.
 XX Morris DW, Engelhard EK;
 XX WPI; 2003-513603/48.
 XX New recombinant nucleic acid comprising a nucleotide sequence of any of
 PT the carcinoma-associated (CA) genes, useful for screening for drug
 PT candidates for diagnosing or treating carcinomas.
 XX
 XX Claim 1; SEQ ID NO 234; 983pp; English.
 PS

XX The invention relates to a recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the fully defined carcinoma-
 CC associated (CA) genes from the 50 tables given in the specification. The
 CC CA proteins are secreted, transmembrane or intracellular proteins. The
 CC recombinant nucleic acids are useful for screening for drug candidates
 CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
 CC ADC8514 represent CA genes of the invention.

XX SQ Sequence 1530 BP; 352 A; 438 C; 451 G; 289 T; 0 U; 0 Other;
 Query Match 1.7%; Score 20; DB 9; Length 1530;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGAGC 766
 |||||
 DB 449 CCTTCCTCATCCGGGAGAGC 468

RESULT 59
 ADC99100
 ID ADC99100 standard; cDNA; 1928 BP.
 AC ADC99100;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human KPP CDNA - SEQ ID 53.
 XX
 KW anti-HIV; anti-allergic; anti-inflammatory; antianaemic; antiparkinsonian;
 KW neurotropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;
 KW immunosuppressive; antichryoid; cytostatic; hepatotropic; dermatological;
 KW antidiabetic; nephrotropic; antigout; thyromimetic; neuroprotective;
 KW osteopathic; antiarthritic; antiparasitic; antihelminthic; antipsoriatic;
 KW uropathic; opthamological; antirheumatic; haemostatic; antibacterial;
 KW virucide; protozoacide; fungicide; kinase; phosphatase; KPP;
 KW cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;
 KW cancer; developmental; mental retardation; neurological;
 KW Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;
 KW diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;
 KW helminthic infection; transgenic; gene therapy; human; ss; gene.
 XX
 OS Homo sapiens.
 XX
 FN WO2003033680-A2.
 XX
 PD 24-APR-2003.
 XX
 PF 17-OCT-2002; 2002WO-US033723.
 XX
 PR 19-OCT-2001; 2001US-0345474P.
 PR 02-NOV-2001; 2001US-0343910P.
 PR 13-NOV-2001; 2001US-0333098P.
 PR 16-NOV-2001; 2001US-0332424P.
 PR 30-NOV-2001; 2001US-0334288P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Bandman O, Baughn MR, Becha SD, Borowsky ML, Duggan BM;
 PI Emerling BM, Forsythe IJ, Gandhi AR, Gorvad AE, Griffin JA;
 PI Gururajan R, Hafalia AJA, Khan FA, Lal PG, Lee EA, Lee SY;
 PI Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;
 PI Ramkumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT;
 PI Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Yao MG, Yue H;
 PI Zebbarjadian Y;
 XX
 DR WPI; 2003-403214/38.
 DR P-PSDB; ADC99048.
 XX
 XX New human kinases and phosphatases and polynucleotides, useful for
 PT diagnosing, treating or preventing autoimmune or inflammatory disorders
 PT (e.g. AIDS, allergy or anemia); multiple sclerosis, osteoarthritis,

PT cancer or hepatitis.
 XX
 PS Claim 5; SEQ ID NO 53; 424pp; English.
 XX
 CC The invention relates to a novel isolated polypeptide which is a human
 CC kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,
 CC agonists and antagonists are useful for diagnosing, treating or
 CC preventing cell proliferative disorders such as atherosclerosis,
 CC cirrhosis, hepatitis and cancer, developmental disorders e.g. mental
 CC retardation, neurological disorders including Alzheimer's disease and
 CC Parkinson's disease, autoimmune and inflammatory disorders such as
 CC Crohn's disease and diabetes mellitus and finally, viral, bacterial,
 CC fungal, parasitic, protozoan or helminthic infections. Furthermore, the
 CC polynucleotides encoding KPP may be useful for creating transgenic
 CC animals to model human disease, as well as during gene therapy
 CC procedures. The current sequence is that of the human KPP cDNA of the
 CC invention.

XX SQ Sequence 1928 BP; 436 A; 555 C; 540 G; 397 T; 0 U; 0 Other;
 Query Match 1.7%; Score 20; DB 9; Length 1928;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGAGC 766
 |||||
 DB 562 CCTTCCTCATCCGGGAGAGC 581

RESULT 60
 AAZ46491
 ID AAZ46491 standard; DNA; 2032 BP.
 XX
 AC AAZ46491;
 XX
 DT 13-MAR-2000 (first entry)
 XX
 DE PKA substrate, Src-family protein encoding DNA.
 XX
 KW Protein kinase A; PKA; PKA signaling pathway; phosphorylation; cancer;
 KW kinase substrate; immunosuppressive disorder; proliferative disease;
 KW HIV infection; AIDS; immunodeficiency; autoimmune disease;
 KW systemic lupus erythematosus; Src-family; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 52..1581
 FT /*tag= a
 FT
 XX
 XX WO9962315-A2.
 XX
 PD 02-DEC-1999.
 XX
 XX 27-MAY-1999; 99WO-GB001680.
 XX
 PR 27-MAY-1998; 98NO-00002419.
 PR 30-DEC-1998; 98US-0114240P.
 XX
 XX (LAUR-) LAURAS AS.
 XX (JONE/) JONES E L.
 PA
 PI Hansson V, Levy FO, Mustelin T, Skalhogg BS, Sundvold V;
 PI Tasken K, Vang T, Altman A, Munshi A;
 XX
 XX WPI; 2000-086801/07.
 DR P-PSDB; AAY49420.
 XX
 XX Altering the activity of protein kinase signaling pathways, used for
 PT treating immunosuppressive disorders, e.g. AIDS, proliferative disorders,
 PT e.g. cancers or autoimmune diseases.
 XX
 PS Claim 22; Page 94-95; 111pp; English.

XX The invention provides a novel method of altering the activity of the
 CC protein kinase A (PKA) signaling pathway in a cell that comprises
 CC altering the extent of phosphorylation of one or more PKA substrates, or
 CC kinase substrates downstream in the PKA signaling pathway. Pharmaceutical
 CC compositions containing a nucleic acid molecule that encodes a PKA
 CC substrate, or fragment, precursor or functionally equivalent variant,
 CC where the sequence is modified to alter its susceptibility to
 CC phosphorylation by PKA can be used for treating a disorder exhibiting
 CC abnormal PKA signaling activity, immunosuppressive disorders or
 CC proliferative diseases. They can be used for treating e.g. HIV infection,
 CC AIDS, common variable immunodeficiency or cancers. Conditions in which
 CC upregulation of the PKA pathway is required, such as autoimmune disease,
 CC e.g. systemic lupus erythematosus, may also be treated. The present
 CC sequence represents a DNA sequence encoding a PKA substrate, wherein the
 CC substrate is in the Src-family, preferably Lck, Fyn, Src, Yes, Fgr, Lyn,
 CC Hck Btk, Yrk, C-tkl, Fyk, Src-1 or Src-2
 XX

SQ Sequence 2032 BP; 450 A; 576 C; 584 G; 422 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 3; Length 2032;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGAGC 766
 |||||
 Db 500 CCTTCTCATCCGGGAGAGC 519

RESULT 61

ACC72850
 ID ACC72850 standard; cDNA; 2032 BP.
 XX
 AC ACC72850;
 XX
 DT 09-JUL-2003 (first entry)
 XX
 DE Human cancer related protein encoding cDNA SEQ ID NO:188.
 XX
 KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
 KW heart disease; atherosclerosis; endometriosis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003025138-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 17-SEP-2002; 2002WO-US029560.
 XX
 PR 17-SEP-2001; 2001US-0323469P.
 PR 20-SEP-2001; 2001US-0323887P.
 PR 13-NOV-2001; 2001US-0350866P.
 PR 08-FEB-2002; 2002US-0355145P.
 PR 08-FEB-2002; 2002US-0355257P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (ECSB-) BOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
 PI Zlotnik A;
 XX
 DR WPI; 2003-354600/33.
 DR P-PSDB; ABR56599.
 XX
 XX New genes that are up-regulated or down-regulated in cancers, useful as
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 PT therapeutic targets for screening drugs for treating these diseases.
 XX
 XX Claim 8; Page 729; 767pp; English.
 XX
 XX The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-

CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
 CC related gene nucleotide sequences which encode the proteins given in
 CC ABR58521 to ABR58709. Also described: (1) determining the presence or
 CC absence of a pathological cell in a patient; (2) an expression vector
 CC comprising a nucleic acid molecule described above; (3) a host cell
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
 CC of (4); (6) specifically targeting a compound to a pathological cell in a
 CC patient by administering to the patient the antibody above; and (7) a
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or
 CC therapeutic targets. In particular, the nucleic acid is useful for
 CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
 CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
 CC atherosclerosis and endometriosis. The nucleic acid is also useful in
 CC drug screening, particularly for identifying agents for treating these
 CC pathologies
 XX

SQ Sequence 2032 BP; 450 A; 579 C; 581 G; 422 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 7; Length 2032;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGAGC 766
 |||||
 Db 500 CCTTCTCATCCGGGAGAGC 519

RESULT 62

ADA02967
 ID ADA02967 standard; cDNA; 2032 BP.
 XX
 AC ADA02967;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human LCK carcinoma associated cDNA, SEQ ID NO:1485.
 XX
 KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003057146-A2.
 XX
 PD 17-JUL-2003.
 XX
 PF 26-DEC-2002; 2002WO-US041414.
 XX
 PR 26-DEC-2001; 2001US-00035832.
 XX
 PA (SAGR-) SAGRES DISCOVERY.
 XX
 PI Morris DW;
 XX
 PI WPI; 2003-587068/55.
 XX
 PT New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 XX
 PS Claim 1; SEQ ID NO 1485; 245pp; English.
 XX
 XX The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism

CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed human CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2032 BP; 450 A; 576 C; 584 G; 422 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 8; Length 2032;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTTCCTCATCCGGGAGAGC 766

Db 500 CCTTCCTCATCCGGGAGAGC 519

RESULT 63

ADB72705

ID ADB72705 standard; mRNA; 2032 BP.

XX AC ADB72705;

XX DT 04-DEC-2003 (first entry)

XX XX Human LCK mRNA.

XX DE human; ss; cytotstatic; gene therapy; vaccine; carcinoma; lymphomas;

XX KW cancer; neoplasm; adenocarcinoma; sarcoma.

XX XX Homo sapiens.

XX XX WO2003008583-A2.

XX XX 30-JAN-2003.

XX XX 26-DEC-2001; 2001WO-US051291.

XX XX 02-MAR-2001; 2001US-00798596.

XX XX 23-OCT-2001; 2001US-00004113.

XX XX 08-NOV-2001; 2001US-00052482.

XX XX 30-NOV-2001; 2001US-00997722.

XX XX 20-DEC-2001; 2001US-00034650.

XX XX (SAGR-) SAGRES DISCOVERY.

XX XX Morris DW, Engelhard EK;

XX XX WPI; 2003-239337/23.

XX XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,

XX XX cancers, neoplasm, adenocarcinoma, or sarcomas.

XX XX Claim 1; SEQ ID NO 533; 2304pp; English.

XX XX The invention relates to a novel recombinant nucleic acid comprising a

XX XX nucleotide sequence selected from any of the 660 sequences fully defined

XX XX in the specification. A polynucleotide of the invention has cytostatic

XX XX activity, and may have a use in gene therapy, or in a vaccine. The

XX XX recombinant nucleic acids and polypeptides are useful for treating

XX XX carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and

XX XX sarcomas. The present sequence represents a human mRNA of the invention.

XX XX Sequence 2032 BP; 450 A; 576 C; 584 G; 422 T; 0 U; 0 Other;

Query Match

Best Local Similarity

Matches 20; Conservative

Qy 747 CCTTCCTCATCCGGGAGAGC 766

Db 500 CCTTCCTCATCCGGGAGAGC 519

RESULT 64

ADC85447

ID ADC85447 standard; DNA; 2032 BP.

XX XX ADC85447;

XX DT 01-JAN-2004 (first entry)

XX DE Human Lck mRNA sequence.

XX KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;

XX KW secreted; transmembrane; intracellular; ds.

XX OS Homo sapiens.

XX XX WO2003045230-A2.

XX XX 05-JUN-2003.

XX XX 02-DEC-2002; 2002WO-US038582.

XX XX 30-NOV-2001; 2001US-00997722.

XX XX (SAGR-) SAGRES DISCOVERY.

XX XX Morris DW, Engelhard EK;

XX XX WPI; 2003-513603/48.

XX XX New recombinant nucleic acid comprising a nucleotide sequence of any of

XX XX the carcinoma-associated (CA) genes, useful for screening for drug

XX XX candidates for diagnosing or treating carcinomas.

XX XX Claim 1; SEQ ID NO 233; 983pp; English.

XX XX The invention relates to a recombinant nucleic acid comprising a

XX XX nucleotide sequence selected from any of the fully defined carcinoma-

XX XX associated (CA) genes from the 50 tables given in the specification. The

XX XX CA proteins are secreted, transmembrane or intracellular proteins. The

XX XX recombinant nucleic acids are useful for screening for drug candidates

XX XX for diagnosing or treating carcinomas. Sequences given in ADC85215-

XX XX ADC85514 represent CA genes of the invention.

XX XX Sequence 2032 BP; 450 A; 576 C; 584 G; 422 T; 0 U; 0 Other;

Query Match

Best Local Similarity

Matches 20; Conservative

Qy 747 CCTTCCTCATCCGGGAGAGC 766

Db 500 CCTTCCTCATCCGGGAGAGC 519

RESULT 65

ADE40448

ID ADE40448 standard; DNA; 2032 BP.

XX XX ADE40448;

XX AC ADE40448;

XX DT 29-JAN-2004 (first entry)

XX XX Human proto-oncogene Tyr protein kinase LCK (gene ID 1611) DNA.

XX AIDS; acquired immunodeficiency syndrome; human immunodeficiency virus;
 KW HIV-related disorder; differential expression; drug screening;
 KW viral replication modulation; diagnosis; prognosis; predisposition;
 KW anti-HIV; gene therapy; antisense therapy; human;
 KW proto-oncogene Tyr protein kinase LCK; gene; ds.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 52..1581
 FT /tag= a
 FT /product= "Human proto-oncogene Tyr protein kinase LCK"
 PN WO2003070883-A2.
 XX
 XX
 PD 28-AUG-2003.
 XX
 XX 13-FEB-2003; 2003WO-US004246.
 XX
 XX 15-FEB-2002; 2002US-0357391P.
 PR 13-MAY-2002; 2002US-0380249P.
 PR 25-JUN-2002; 2002US-0391306P.
 PR 27-AUG-2002; 2002US-0405297P.
 PR 19-SEP-2002; 2002US-0412007P.
 PR 10-OCT-2002; 2002US-0417508P.
 PR 10-DEC-2002; 2002US-0432318P.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 XX Powell DM, Weich NS;
 FI
 XX WPI; 2003-671808/63.
 DR P-PSDB; ADE40449.
 DR
 XX
 XX Identifying a compound capable of diagnosing, preventing or treating AIDS
 PT or an HIV-related disorder comprises assaying the ability of the compound
 PT to modulate e.g. 1414, 1481 or 1553 nucleic acid expression or
 PT polypeptide activity.
 XX
 XX Claim 1; SEQ ID NO 27; 167pp; English.
 PS
 XX The invention relates to a method of identifying a compound useful in the
 CC treatment of AIDS (acquired immunodeficiency syndrome) or an HIV (human
 CC immunodeficiency virus)-related disorder. The invention involves assaying
 CC the ability of a test compound to modulate the activity or expression of
 CC 26 human proteins. These proteins and nucleic acids encoding them
 CC (ADE40422-ADE40473) are differentially expressed in tissues relating to
 CC AIDS or an HIV-related disorder compared to their expression in normal
 CC tissues. The invention also relates to the use of the compounds
 CC identified to modulate viral replication in a cell and to treat a patient
 CC with AIDS or an HIV-related disorder. The invention further discloses
 CC methods for the diagnostic evaluation and prognosis of various HIV-
 CC related disorders, and for the identification of individuals exhibiting a
 CC predisposition to such conditions. The modulatory compounds identified
 CC using the method of the invention may be small organic molecules,
 CC peptides, antibodies or antisense nucleic acid molecules. The methods of
 CC the invention are useful in diagnosing, preventing or treating AIDS or
 CC HIV-related disorders. The present sequence represents a human
 CC polynucleotide which is differentially expressed in AIDS or HIV-related
 CC disorders.
 XX
 XX Sequence 2032 BP; 450 A; 579 C; 581 G; 422 T; 0 U; 0 Other;
 SQ
 Query Match 1.7%; Score 20; DB 9; Length 2032;
 Best Local Similarity 100.0%; Fred. NO. 46;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 747 CCTTCTCATCCGGAGAGC 766
 DB 500 CCTTCTCATCCGGAGAGC 519

RESULT 66
 ABS65202
 ID ABS65202 standard; cDNA; 2034 BP.
 XX
 AC ABS65202;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE cDNA encoding tumour involved gene (TIG) splice variant, NV-3.
 XX
 KW Human; ss; gene; splice variant; tumour-involved gene; TIG;
 KW pharmaceutical compound; cancer; diagnostic; tumour; gene therapy;
 KW endothelial cell; cell differentiation; cell proliferation; apoptosis;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 EN US2002086384-A1.
 XX
 PD 04-JUL-2002.
 XX
 XX 13-MAR-2001; 2001US-00805020.
 XX
 XX 14-MAR-2000; 2000IL-00135402.
 PR 16-MAY-2000; 2000IL-00136154.
 XX
 XX (LEVI/) LEVINE Z.
 PA (DAVI/) DAVID A.
 PA (ROMA/) ROMANO C.
 PA (BERN/) BERNSTEIN J.
 XX
 XX Levine Z, David A, Romano C, Bernstein J;
 PI
 XX WPI; 2002-635679/68.
 DR P-PSDB; ABG79672.
 DR
 XX Novel nucleic acid sequence, which is an alternative splicing variant of
 PT tumor involved genes, useful for detecting cancer, predisposition to
 PT cancer, for evaluating cancer state and in gene therapy for treating
 PT cancer.
 XX
 XX Claim 1; Page 27-28; 180pp; English.
 PS
 XX The invention discloses isolated human nucleic acid alternative splicing
 CC variants that are all tumour-involved genes (TIGs). The nucleic acids and
 CC polypeptides are useful for determining the level of a nucleic acid or
 CC polypeptide in a biological sample, for detecting a variant nucleic acid
 CC or polypeptide sequence in a biological sample, for determining the level
 CC of variant nucleic acid or polypeptide sequences in a biological sample
 CC and for determining the ratio between the level of variant sequence in a
 CC first biological sample and the level of the original sequence from which
 CC the variant has been varied by alternative splicing in a second
 CC biological sample and for raising antibodies. A pharmaceutical
 CC composition comprising a carrier and the nucleic acid, is useful for
 CC treating diseases (e.g. cancer) that can be ameliorated or cured by
 CC increasing or decreasing the level of the encoded protein. The nucleic
 CC acids are also useful for diagnostic purposes, especially for detecting
 CC cancer or a predisposition to cancer, for evaluating the state or
 CC aggressiveness of cancer disease, in basic research, for understanding
 CC the physiological function of the original TIG, in targeting or
 CC developing pharmaceuticals, for distinguishing various stages in the life
 CC cycle of the same type of cells which may be helpful for the development
 CC of pharmaceuticals for various cancer stages in which cell cycle is non-
 CC normal, for determining mutations in tumour-involved genes and in gene
 CC therapy. The polypeptides are useful for identifying compounds capable of
 CC binding to the variant product and modulating its activity and for
 CC modulating endothelial differentiation and proliferation, as well as to
 CC modulate apoptosis either ex vivo or in vivo. The sequences presented in
 CC ABS65200-ABS65235 are the coding sequences for the new variants (NV) 1-
 CC 36 of the TIGs disclosed
 XX
 XX Sequence 2034 BP; 455 A; 570 C; 586 G; 422 T; 0 U; 1 Other;
 SQ

Query Match 1.7%; Score 20; DB 6; Length 2034;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CTTCTCTCATCCGGGAGGC 766
|||||
DB 562 CTTCTCTCATCCGGGAGGC 581
|||||

RESULT 67
ACA56854
ID ACA56854 standard; cDNA; 2129 BP.
XX
AC ACA56854;
XX
DT 06-JUN-2003 (first entry)
XX
DE Human signalling pathway polynucleotide probe SEQ ID NO 1452.
XX
KW Human; probe; ss; array element; Parkinson's disease;
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS Homo sapiens.
XX
PN US6500938-B1.
XX
PD 31-DEC-2002.
XX
PF 30-JAN-1998; 98US-00016434.
XX
PR 30-JAN-1998; 98US-00016434.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Seilhamer JJ;
XX
DR WPI; 2003-352189/33.
XX
PT Combination of polynucleotide probes, useful as array elements in a
PT microarray for monitoring the expression of a number of target
PT polynucleotides.
XX
PS Claim 1; SEQ ID NO 1452; 65pp; English.
XX
CC The invention relates to a combination which, comprises a number of
CC polynucleotide probes comprising a sequence selected from one of the 1490
CC sequences mentioned in the specification. The combination is useful as an
CC array element in a microarray for monitoring the expression of a number
CC of target polynucleotides. The microarray is particularly useful in the
CC diagnosis and treatment of cancer and immunopathology and neuropathology.
CC The microarray is useful in diagnostics and treatment regimens, drug
CC discovery and development, toxicological and carcinogenicity studies,
CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, CDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signaling pathway populations which can be used to diagnose
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC and Parkinson's disease. The present sequence represents a polynucleotide
CC probe of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=06500938B1
XX
SQ Sequence 2129 BP; 474 A; 593 C; 623 G; 439 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 7; Length 2129;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CTTCTCTCATCCGGGAGGC 766
|||||
DB 508 CTTCTCTCATCCGGGAGGC 527
|||||

RESULT 68
ACC81082
ID ACC81082 standard; mRNA; 2129 BP.
XX
AC ACC81082;
XX
DT 25-JUL-2003 (first entry)
XX
DE Human T-lymphocyte specific protein tyrosine kinase p56lck aberrant mRNA.
XX
KW Human; T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; asthma;
KW immunosuppressive; antiasthmatic; antiallergic; antiinflammatory;
KW lymphocyte activation; lymphocyte migration; cytokine production;
KW cell surface marker expression; antibody production; apoptosis; allergy;
KW antibody proliferation; antibody differentiation; hypersensitivity;
KW graft versus host disease; inflammation; gene; ss; T-lymphocyte;
KW protein tyrosine kinase; p56lck.
XX
OS Homo sapiens.
XX
PN WO2003029277-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031618.
XX
PR 03-OCT-2001; 2001US-0327212P.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Chu P, Li C, Liao XC, Masuda E, Pardo J, Zhao H;
XX
DR WPI; 2003-363276/34.
XX
PT P-PSDB; ABR59690.
XX
PT Identifying a compound that modulates T lymphocyte activation, useful for
PT monitoring changes in cell surface marker expression, comprises
PT contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with
PT a compound.
XX
PS Disclosure; Page 63-64; 126pp; English.
XX
CC The invention relates to a novel method for identifying a compound that
CC modulates T lymphocyte activation. The method comprises contacting a T
CC cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,
CC where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic
CC acid that hybridises to a nucleic acid encoding a polypeptide having a
CC sequence selected from two 606-amino acid sequence and a 415-amino acid
CC sequence given in the specification. The method of the invention has
CC immunosuppressive, antiasthmatic, antiallergic, and antiinflammatory
CC activity. The method is useful for identifying compounds that modulate
CC lymphocyte activation and migration, and for monitoring changes in cell
CC surface marker expression, cytokine production, antibody production,
CC proliferation and differentiation, and apoptosis, using either cell lines
CC or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as
CC drug targets for compounds that suppress or activate lymphocyte
CC activation and migration, e.g. for the treatment of diseases in which
CC modulation of the immune response is desired such as delayed type
CC hypersensitivity reactions, asthma, allergies, graft versus host disease,
CC and acute and chronic inflammation. Modulators of lymphocyte activation
CC are useful for treating disorders related T and B cell activation and
CC migration. The present sequence is used in the exemplification of the
CC invention
XX
SQ Sequence 2129 BP; 474 A; 593 C; 623 G; 439 T; 0 U; 0 Other;

		Best Local Similarity	100.0%; Pred. No. 46;			Mismatches	0; Indels	0; Gaps	0;
		Matches	20; Conservative	0; Mismatches	0; Indels	0; Gaps	0;		
QY	747	CCTTCCTCATCGGGAGAGC	766						
DB	508	CCTTCCTCATCGGGAGAGC	527						
RESULT 69									
ID	ACA64888	standard; DNA; 2129 BP.							
AC	ACA64888;								
XX	XX	(first entry)							
DT	27-JUN-2003								
XX	XX								
DE	Human p56lck DNA corresponding to U23852.								
XX	XX								
KW	Human; chronic inflammatory joint disease; infection; tumour;								
KW	antifibrotic; cytotaxtic; antiarthritic; antirheumatic;								
KW	immunosuppressive; gene therapy; etiological pathogenicity; ds.								
OS	Homo sapiens.								
PA	DE10127572-A1.								
PN	PN								
XX	XX								
PD	05-DEC-2002.								
PF	30-MAY-2001; 2001DE-01027572.								
XX	XX								
PR	30-MAY-2001; 2001DE-01027572.								
XX	XX								
PA	(PATH-) PATHOARRAY GNBH.								
PI	Haeupl T, Ungethuem U, Blaess S;								
PPT	WPI; 2003-240797/24.								
PT	Reagents for diagnosis, study and therapy of chronic inflammatory joint								
PT	and other diseases, comprises any of many specified genes or derived								
PT	proteins.								
PS	Claim 1; Page; 12pp; German.								
XX	XX								
CC	This invention describes a novel reagent for diagnosis, molecular								
CC	definition and therapy of chronic inflammatory joint diseases, and other								
CC	inflammatory disorders, infective or tumour diseases in humans. The								
CC	products of the invention have anti-inflammatory, cytostatic,								
CC	antiarthritic, antirheumatic and immunosuppressive activity and can be								
CC	used for gene therapy. The reagent of the invention and any proteins and								
CC	antibodies derived from it, are used (i) for analysing tissue and blood								
CC	samples for medical diagnosis; (ii) for diagnosis and characterisation of								
CC	chronic joint diseases, on the basis of molecular characterisation, and								
CC	determining the etiological pathogenicity principle of as yet								
CC	uncharacterised inflammatory diseases, also monitoring progression and/or								
CC	treatment of disease, and optimisation of therapy and (iii) for								
CC	developing treatments for inflammatory diseases, particularly of joints,								
CC	infections and tumours. ACA64801-ACA64965 represent human polynucleotides								
CC	used in the method of the invention								
XX	Sequence 2129 BP; 474 A; 593 C; 623 G; 439 T; 0 U; 0 Other;								
SQ									
Query Match		1.7%; Score 20; DB 7; Length 2129;							
Best Local Similarity		100.0%; Pred. No. 46;							
Matches	20; Conservative	0; Mismatches	0; Indels	0; Gaps	0;				
QY	747	CCTTCCTCATCGGGAGAGC	766						
DB	508	CCTTCCTCATCGGGAGAGC	527						
RESULT 70									
ID	ABSG5203	standard; cDNA; 2282 BP.							
AC	ABSG5203;								
XX	XX								
DT	15-NOV-2002	(first entry)							
XX	XX								
DE	cDNA encoding tumour involved gene (TIG) splice variant, NV-4.								
XX	XX								
KW	Human; ss; gene; splice variant; tumour-involved gene; TIG;								
KW	pharmaceutical composition; cancer; diagnostic; tumour; gene therapy;				</				

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGGC 766
 |||||
 Db 736 CCTTCCTCATCCGGGAGGC 755

RESULT 71
 AAS86451
 ID AAS86451 standard; cDNA; 2320 BP.
 AC AAS86451;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #22255.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 XX
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG22264.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 22255; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcr_sequences
 XX
 SQ Sequence 2320 BP; 529 A; 649 C; 659 G; 483 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 5; Length 2320;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGGC 766
 |||||
 Db 734 CCTTCCTCATCCGGGAGGC 753

RESULT 72
 AAH46496/C
 ID AAH46496 standard; cDNA; 2483 BP.
 AC AAH46496;
 XX
 DT 17-SEP-2001 (first entry)
 XX
 DE Subtilopeptidase 9 coding sequence.
 XX
 KW Subtilopeptidase 9; cytostatic; haemostatic; virucide; immunomodulatory;
 KW antiinflammatory; gene therapy; malignant tumour; haemopathy;
 KW HIV infection; immunological disease; inflammation; ss.
 XX
 OS Unidentified.
 XX
 FN WO200146247-A1.
 XX
 PD 28-JUN-2001.
 XX
 PF 11-DEC-2000; 2000WO-CN000559.
 XX
 PR 21-DEC-1999; 99CN-00124299.
 XX
 PA (UVFU-) UNIV FUDAN.
 PA (SHAN-) SHANGHAI BIO DOOR GENE TECHNOLOGY LTD.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI; 2001-418036/44.
 DR P-PSDB; AAG64202.
 XX
 PT Subtilopeptidase 9 and encoded polynucleotide, used in diagnosis and
 PT treatment of malignant tumors, hemopathy, human immunodeficiency virus
 PT infection, immunological diseases and inflammation.
 XX
 PS Claim 6; Page 28-29; 34pp; Chinese.
 XX
 CC The present sequence is the coding sequence for subtilopeptidase 9. The
 CC subtilopeptidase and coding sequence are useful in the diagnosis and
 CC treatment of malignant tumour, haemopathy, human immunodeficiency virus
 CC (HIV) infection, immunological diseases and various inflammations
 XX
 SQ Sequence 2483 BP; 823 A; 374 C; 430 G; 856 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 4; Length 2483;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 TGTGACCATGGAGACAGAGA 491
 |||||
 Db 971 TGTGACCATGGAGACAGAGA 952

RESULT 73
 ADA02966
 ID ADA02966 standard; DNA; 31842 BP.
 AC ADA02966;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human LCK carcinoma associated gene, SEQ ID NO:1484.
 XX
 KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW gene; ds.
 XX

OS Homo sapiens.
 PN WO2003057146-A2.
 XX
 PD 17-JUL-2003.
 XX
 PF 26-DEC-2002; 2002WO-US041414.
 XX
 XX 26-DEC-2001; 2001US-00035832.
 XX
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 XX Morris DW;
 XX
 XX WPI; 2003-587068/55.
 DR
 XX New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 XX
 XX Claim 1; SEQ ID NO 1484; 245pp; English.
 PS
 XX The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed human CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 31842 BP; 6799 A; 6356 C; 6981 G; 7010 T; 0 U; 4696 Other;
 Query Match 1.7%; Score 20; DB 8; Length 31842;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 747 CCTTCTCATCCGGGAGAGC 766
 Db 11316 CCTTCTCATCCGGGAGAGC 11335
 RESULT 74
 ID ADB72704 standard; DNA; 31842 BP.
 AC ADB72704;
 XX
 DT 04-DEC-2003 (first entry)
 DE Human LCK gene.
 XX
 KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
 XX
 OS Homo sapiens.
 XX
 PN WO2003008583-A2.
 XX
 XX 30-JAN-2003.

XX 26-DEC-2001; 2001WO-US051291.
 PF
 XX 02-MAR-2001; 2001US-00798586.
 PR
 XX 23-OCT-2001; 2001US-00004113.
 PR
 XX 08-NOV-2001; 2001US-00052482.
 PR
 XX 30-NOV-2001; 2001US-00997722.
 PR
 XX 20-DEC-2001; 2001US-00034650.
 XX
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 XX Morris DW, Engelhard EK;
 XX
 XX WPI; 2003-239337/23.
 DR
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX
 XX Claim 1; SEQ ID NO 532; 2304pp; English.
 PS
 XX The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined
 CC in the specification. A polynucleotide of the invention has cytostatic
 CC activity, and may have a use in gene therapy, or in a vaccine. The
 CC recombinant nucleic acids and polypeptides are useful for treating
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 CC sarcomas. The present sequence represents a human gene of the invention.
 XX
 SQ Sequence 31842 BP; 6799 A; 6356 C; 6981 G; 7010 T; 0 U; 4696 Other;
 Query Match 1.7%; Score 20; DB 9; Length 31842;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 747 CCTTCTCATCCGGGAGAGC 766
 Db 11316 CCTTCTCATCCGGGAGAGC 11335
 RESULT 75
 ID ADB85446
 XX ADB85446 standard; DNA; 31842 BP.
 AC ADB85446;
 XX
 DT 01-JAN-2004 (first entry)
 DE Human Lck genomic sequence.
 XX
 KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
 KW secreted; transmembrane; intracellular; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003045230-A2.
 XX
 PD 05-JUN-2003.
 XX
 XX 02-DEC-2002; 2002WO-US038582.
 PF
 XX 30-NOV-2001; 2001US-00997722.
 PR
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 XX Morris DW, Engelhard EK;
 XX
 XX WPI; 2003-513603/48.
 DR
 XX New recombinant nucleic acid comprising a nucleotide sequence of any of
 PT the carcinoma-associated (CA) genes, useful for screening for drug
 PT candidates for diagnosing or treating carcinomas.
 XX
 PS Claim 1; SEQ ID NO 232; 983pp; English.

XX The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carcinoma-
CC associated (CA) genes from the 50 tables given in the specification. The
CC CA proteins are secreted, transmembrane or intracellular proteins. The
CC recombinant nucleic acids are useful for screening for drug candidates
CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC ADC85514 represent CA genes of the invention.
XX
SQ Sequence 31842 BP; 6799 A; 6356 C; 6981 G; 7010 T; 0 U; 4596 Other;

Query Match 1.7%; Score 20; DB 9; Length 31842;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGAC 766
DB 11316 CCTTCTCATCCGGGAGAC 11335

RESULT 76
ABQ74964_4
Continuation (5 of 8) of ABQ74964 from base 400001 (Human kinase protein genomic DNA seq
WP Sequence split into 8 fragments LOCUS ABQ74964 Accession Abq74964
WP Fragment Name Begin End
WP ABQ74964_0 1 110000
WP ABQ74964_1 100001 210000
WP ABQ74964_2 200001 310000
WP ABQ74964_3 300001 410000
WP ABQ74964_4 400001 510000
WP ABQ74964_5 500001 610000
WP ABQ74964_6 600001 710000
WP ABQ74964_7 700001 786431

Query Match 1.7%; Score 20; DB 6; Length 110000;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 TGAGCAGGGAGAAAGCAGAG 712
DB 12751 TGAGCAGGGAGAAAGCAGAG 12770

RESULT 77
ABN31762
ID ABN31762 standard; DNA; 65 BP.
XX
AC ABN31762;
XX
DT 15-JUL-2002 (first entry)
XX
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:4510.
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Rattus norvegicus.
XX
FN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB001903.
XX
PR 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
XX
PA (COMP-) COMPUGEN INC.
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX

PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
PS Example 1; SEQ ID NO 4510; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridizing selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 65 BP; 16 A; 13 C; 18 G; 18 T; 0 U; 0 Other;

Query Match 1.6%; Score 19; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 395 AACATGGGAAGTCTGCCCA 413
DB 2 AACATGGGAAGTCTGCCCA 20

RESULT 78
AAK24414/C
ID AAK24414 standard; DNA; 114 BP.
XX
AC AAK24414;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 24405.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
FN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 XX brains.
 XX Example 4; SEQ ID NO 24405; 650pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 XX probes which are derived from genomic sequences expressed in the human
 XX brain. They can be used to measure gene expression in brain cell samples,
 XX which may enable the diagnosis and improved treatment of nervous system
 XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 XX epilepsy and cancers. The present sequence is one of the probes of the
 XX invention
 XX Sequence 114 BP; 24 A; 35 C; 30 G; 25 T; 0 U; 0 Other;
 XX
 XX Query Match 1.6%; Score 19; DB 4; Length 114;
 XX Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 751 CCTCATCCGGGAGAGCCAG 769
 XX 51 CCTCATCCGGGAGAGCCAG 33
 XX
 XX RESULT 79
 XX ACL21602
 XX ID ACL21602 standard; DNA; 307 BP.
 XX
 XX AC ACL21602;
 XX
 XX 27-OCT-2003 (revised)
 XX DT 17-OCT-2003 (first entry)
 XX
 XX DNA clone originating in barley containing SNP encoding sequence #11593.
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 XX gene, ss.
 XX Hordeum vulgare; var. (cul.Akashinriki).
 XX WO2003057877-A1.
 XX 17-JUL-2003.
 XX
 XX 16-DEC-2002; 2002WO-IB005403.
 XX
 XX 20-DEC-2001; 2001JP-00387059.
 XX 20-DEC-2001; 2001JP-00387131.
 XX 20-DEC-2001; 2001JP-00403299.
 XX 20-DEC-2001; 2001JP-00403300.
 XX 27-SEP-2002; 2002JP-00327515.
 XX
 XX (UYN1-) UNIV JAPAN OKAYAMA.
 XX
 XX Sato K, Takeda K, Kohara Y;
 XX WPI; 2003-587127/55.
 XX
 XX Single nucleotide polymorphism sites in barley varieties and DNA
 XX sequences containing them for analysis and identification of barley
 XX varieties and production of barley transformants with desired
 XX characteristics.
 XX Disclosure; SEQ ID XX; 284pp; Japanese.
 XX The present invention relates to oligonucleotide clones originating in
 XX barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 XX (SNP). The oligonucleotides may be used for analysis of SNPs among barley

CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)
 XX
 XX SQ Sequence 307 BP; 70 A; 85 C; 90 G; 58 T; 0 U; 4 Other;
 XX
 XX Query Match 1.8%; Score 19; DB 8; Length 307;
 XX Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 303 CAGAGCATGCGTCTCAGCA 321
 XX 106 CAGAGCATGCGTCTCAGCA 124
 XX
 XX RESULT 80
 XX AAK83462/c
 XX ID AAK83462 standard; DNA; 416 BP.
 XX
 XX AC AAK83462;
 XX
 XX DT 07-NOV-2001 (first entry)
 XX
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38274.
 XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX cytostatic; gene therapy; vaccine; metastasis; ds.
 XX Homo sapiens.
 XX WO200157182-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US001354.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 XX 04-FEB-2000; 2000US-0180628P.
 XX 24-FEB-2000; 2000US-0184664P.
 XX 02-MAR-2000; 2000US-0186350P.
 XX 16-MAR-2000; 2000US-0189874P.
 XX 17-MAR-2000; 2000US-0190076P.
 XX 18-APR-2000; 2000US-0198123P.
 XX 19-MAY-2000; 2000US-0205515P.
 XX 07-JUN-2000; 2000US-0209467P.
 XX 30-JUN-2000; 2000US-0215135P.
 XX 07-JUL-2000; 2000US-0216647P.
 XX 07-JUL-2000; 2000US-0216880P.
 XX 11-JUL-2000; 2000US-0217487P.
 XX 11-JUL-2000; 2000US-0217496P.
 XX 14-JUL-2000; 2000US-0218290P.
 XX 26-JUL-2000; 2000US-0220963P.
 XX 26-JUL-2000; 2000US-0220964P.
 XX 14-AUG-2000; 2000US-0224518P.
 XX 14-AUG-2000; 2000US-0224519P.
 XX 14-AUG-2000; 2000US-0225213P.
 XX 14-AUG-2000; 2000US-0225214P.
 XX 14-AUG-2000; 2000US-0225266P.
 XX 14-AUG-2000; 2000US-0225267P.
 XX 14-AUG-2000; 2000US-0225268P.
 XX 14-AUG-2000; 2000US-0225270P.
 XX 14-AUG-2000; 2000US-0225447P.
 XX 14-AUG-2000; 2000US-0225477P.
 XX 14-AUG-2000; 2000US-0225757P.
 XX 14-AUG-2000; 2000US-0225758P.
 XX 14-AUG-2000; 2000US-0225759P.
 XX 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227003P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 38274; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK64702 to AAK64703. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX SQ Sequence 416 BP; 109 A; 108 C; 142 G; 57 T; 0 U; 0 Other;
Query Match 1.6%; Score 19; DB 4; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 885 GCCTCACCTTCCTCCTCACT 903
Db 93 GCCTCACCTTCCTCCTCACT 75

RESULT 81
AAK83463/C
ID AAK83463 standard; DNA; 417 BP.
XX
AC AAK83463;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38275.
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227189P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230433P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244517P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246509P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

```

PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 38275; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 417 BP; 102 A; 112 C; 144 G; 59 T; 0 U; 0 Other;
XX
Query Match 1.6%; Score 19; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 885 GCCTCACCTTCCCTCACT 903
Db 93 GCCTCACCTTCCCTCACT 75
RESULT 82
AAK83459/c
ID AAK83459 standard; DNA; 417 BP.
XX
XX AAK83459;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38271.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217486P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225799P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228224P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.

```

02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and metastasis.

PS Disclosure; SEQ ID NO 38271; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK54951 to AAK64702. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX

Query Match 1.6%; Score 19; DB 4; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.5e-02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 885 GCCTCACCTTCCCTCACT 903
 |||||
 Db 93 GCCTCACCTTCCCTCACT 75

RESULT 83
 ABX04900
 ID ABX04900 standard; DNA; 510 BP.
 XX
 AC ABX04900;
 XX
 DT 17-JAN-2003 (first entry)
 XX
 DE Conus sp conotoxin-associated DNA SEQ ID 237.
 XX
 KW Conotoxin; cone snail; analgesic; voltage-gated ion channel modulator;
 KW ligand-gated ion channel modulator; pain-relief; ds.
 XX
 OS Conus textile.
 XX
 PN WO200264740-A2.
 XX
 PD 22-AUG-2002.
 XX
 PF 11-FEB-2002; 2002WO-US003887.
 XX
 PF 09-FEB-2001; 2001US-0267408P.
 XX
 PR (COGN-) COGNETIX INC.
 PR (UTAH) UNIV UTAH RES FOUND.
 XX
 PI Olivera BM, McIntosh JM, Watkins M, Garrett JE, Cruz LJ;
 PI Grille M, Walker CS, Shetty R, Jones RM, Schoenfeld RM;
 XX
 DR WPI; 2002-706921/76.
 DR P-PSDB; ABG99523.
 XX
 XX New cone snail conotoxin peptides, useful as a pain reliever for
 PT alleviating pain in an individual suffering from pain or who is about to
 PT be subjected to a pain-causing event, or for treating voltage-gated ion
 PT channel disorders.
 XX
 XX Claim 1; Page 197; 305pp; English.
 XX
 CC This invention describes novel conotoxin peptides from the cone snail,
 CC genus Conus which have analgesic activity and can act as a voltage-gated
 CC ion channel modulator or a ligand-gated ion channel modulator. The

conotoxin peptide is useful as a pain-relieving agent for alleviating pain in an individual who is either exhibiting pain or is about to be subjected to a pain-causing event. The conotoxin peptide is also useful for treating or preventing disorders associated with voltage-gated ion channel disorders, ligand-gated ion channel disorders or receptor disorders. The radiolabeled conotoxin peptide is also useful for characterising a new site on these receptors or channels, and for screening and identifying novel small molecules that interact with the above-mentioned channels or receptors, which are monamine transporters. ABX04827-ABX04970 encode the conotoxin proteins described in the disclosure of the invention

Sequence 510 BP; 113 A; 135 C; 124 G; 138 T; 0 U; 0 Other;

Query Match 1.6%; Score 19; DB 6; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GGACAAATGGGAAGTCTGC 410
DB 390 GGACAAATGGGAAGTCTGC 408

RESULT 84

AAK11828/c

ID AAK11828 standard; DNA; 599 BP.

XX AC AAK11828;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 11819.

Human; brain expressed exon; gene expression analysis; probe; microarray; Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer; ss.

OS Homo sapiens.

WO200157275-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US000667.

04-FEB-2000; 2000US-0180312P.

26-MAY-2000; 2000US-0207456P.

30-JUN-2000; 2000US-00608408.

03-AUG-2000; 2000US-00632366.

21-SEP-2000; 2000US-0234687P.

27-SEP-2000; 2000US-0236359P.

04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483446/52.

Single exon nucleic acid probes for analyzing gene expression in human brains.

Example 4; SEQ ID NO 11819; 650pp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention

Sequence 599 BP; 181 A; 154 C; 107 G; 157 T; 0 U; 0 Other;

Query Match 1.6%; Score 19; DB 4; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 751 CCTCATCCGGGAGAGCCAG 769

DB 75 CCTCATCCGGGAGAGCCAG 57

RESULT 85

AAF91867/c

ID AAF91867 standard; cDNA; 1033 BP.

XX AC AAF91867;

DT 22-MAY-2001 (first entry)

DE Human secreted protein-encoding gene 10 cDNA clone HTLIT63, SEQ ID NO:20.

Human; secreted protein; proliferative disorder; cancer; tumour; foetal abnormality; developmental abnormality; haematopoietic disorder; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; inflammation; allergy; neurological disorder; Alzheimer's disease; Parkinson's disease; cognitive disorder; schizophrenia; asthma; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; angioinetic disorder; kidney disorder; gastrointestinal disorder; pregnancy-related disorder; endocrine disorder; infection; wound healing; vulnerability; cell culture; chemotaxis; food additive; binding partner identification; ss.

OS Homo sapiens.

WO200118022-A1.

15-MAR-2001.

31-AUG-2000; 2000WO-US024008.

03-SEP-1999; 99US-0152315P.

03-SEP-1999; 99US-0152317P.

(HUMA-) HUMAN GENOME SCI INC.

Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis CA, Rosen CA; Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW; Moore PA, Shi Y, Wei Y, Florence KA;

WPI; 2001-203081/20.

P-PSDB; AAB87351.

Nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers.

Claim 1; Page 494; 607pp; English.

AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted protein genes, and AAB87342-AAB87413 represent the proteins they encode. AAB87414-AAB87454 represent human secreted protein fragments. The genes and their corresponding secreted proteins are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 52 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,

The invention discloses a method for identifying a test compound which binds to a kinase, or phosphatase, or which modulates the activity of kinase, or phosphatase. The method comprises contacting the kinase with the test compound and detecting the binding of the test compound, or determining the compound which modulates the activity of the kinase. Protein kinases and phosphatases play critical roles in the regulation of biochemical and morphological changes associated with cellular growth and division. Also disclosed are isolated nucleic acid molecules designated kinase and phosphatase nucleic acid molecules which encode protein kinase and protein phosphatase polypeptides, antisense nucleic acid molecules of the above nucleic acid molecules, non-human transgenic animals in which a kinase or phosphatase gene is introduced or disrupted, a diagnostic assay

PR 25-JUN-2002; 2002US-0391324P.
PR 15-JUL-2002; 2002US-0395944P.
PR 22-JUL-2002; 2002US-0397726P.
PR 13-AUG-2002; 2002US-0403046P.
PR 22-AUG-2002; 2002US-0405185P.
PR 27-AUG-2002; 2002US-0406361P.
PR 25-OCT-2002; 2002US-0421195P.
PR 12-NOV-2002; 2002US-0425456P.
PR 19-NOV-2002; 2002US-0427626P.
PR 10-DEC-2002; 2002US-0432122P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Hunter JJ, Macheth KJ, Tsai F, Leseon A, Lightcap ES;
PI Williamson MW, Rudolph-Owen LA;
XX P-PSDB; ADE38351.
XX WPI; 2003-646176/61.
DR Treating subject having tumorigenic disorder or angiogenic disorder
PT caused by aberrant polypeptide e.g., N-formylpeptide receptor or nucleic
PT acid, by administering a modulator.
XX
PS Disclosure; SEQ ID NO 11; 454pp; English.
XX
XX This invention relates to a novel method of treating a human subject
CC having a tumorigenic disorder or angiogenic disorder, caused by aberrant
CC gene expression or activity of an isolated protein, by administering a
CC modulator. The modulator may have cytostatic, antithyroid, antidiabetic
CC or ophthalmological activity. The method is useful for treating a subject
CC having a tumorigenic or angiogenic disorder, in particular for treating
CC cancer (for example breast cancer, colon cancer, lung cancer or prostatic
CC cancer) and, for example, Grave's disease and diabetic retinopathy. The
CC present sequence is a DNA sequence which encodes the novel isolated human
CC protein 3702 of the invention.
XX
SQ Sequence 1467 BP; 250 A; 501 C; 466 G; 250 T; 0 U; 0 Other;
Query Match 1.6%; Score 19; DB 9; Length 1467;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 AGGGGCGCTTCTCATCCGG 760
Db 423 AGGGGCGCTTCTCATCCGG 441
RESULT 88
ACCS7671
ID ACCS7671 standard; cDNA; 1490 BP.
XX
XX ACCS7671;
XX
XX 28-JUL-2003 (first entry)
XX
DE Mouse protein tyrosine kinase polynucleotide.
XX
XX Mouse; protein tyrosine kinase; PTK; enzyme; stent; graft; catheter;
XX smooth muscle cell; gene; ss.
XX
XX Mus musculus.
XX
XX WO2003034938-A2.
XX
XX 01-MAY-2003.
XX
XX 25-OCT-2002; 2002WO-US034344.
XX
XX 25-OCT-2001; 2001US-0343732P.
XX
XX (WISC) WISCONSIN ALUMNI RES FOUND.
XX
XX Wolff MR;
PI

XX WPI; 2003-430375/40.
XX
XX Device for stenting blood vessel, e.g. catheter, has coated, adsorbed or
XX impregnated protein tyrosine kinase inhibitor that inhibits vascular
XX smooth muscle cell proliferation within blood vessel proximal to device.
XX
XX Disclosure; Page 108-110; 110pp; English.
XX
XX The present sequence is a polynucleotide for a murine protein tyrosine
XX kinase (PTK). PTKs catalyse the transfer of the gamma-phosphoryl group
XX from ATP to the tyrosine hydroxyl moiety of a protein substrate. The
XX invention is based on the finding that inhibiting the action of PTKs
XX selectively inhibits the proliferation of vascular smooth muscle cells
XX (VSMCs). The invention is directed to methods of selectively inhibiting
XX the proliferation of VSMCs following vascular injury or surgical
XX intervention, such as percutaneous revascularisation. PTK inhibitors,
XX preferably those that inhibit the Bcr-Abl tyrosine kinase, are coated
XX onto vascular stents, native grafts or prosthetic vascular grafts, to
XX prevent VSMC proliferation selectively, while not adversely affecting the
XX proliferation of endothelial cells. A device for stenting a blood vessel,
XX and a method of preventing restenosis following vascular intervention are
XX claimed
XX
SQ Sequence 1490 BP; 318 A; 434 C; 422 G; 316 T; 0 U; 0 Other;
Query Match 1.6%; Score 19; DB 7; Length 1490;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 AGGGGCGCTTCTCATCCGG 760
Db 434 AGGGGCGCTTCTCATCCGG 452
RESULT 89
AAV81743
ID AAV81743 standard; cDNA; 1548 BP.
XX
XX AAV81743;
XX
XX 10-MAR-1999 (first entry)
XX
XX Human SAD encoding cDNA.
XX
XX PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;
XX type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;
XX neurodegenerative disease; neuronal survival; Alzheimer's disease;
XX Parkinson's disease; Huntington's disease; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 49..1515
XX /*tag= a
XX
XX WO9849317-A2.
XX
XX 05-NOV-1998.
XX
XX 27-APR-1998; 98WO-US008439.
XX
XX 28-APR-1997; 97US-0044428P.
XX 20-MAY-1997; 97US-0047222P.
XX 11-JUN-1997; 97US-0049477P.
XX 11-JUN-1997; 97US-0049756P.
XX 18-JUN-1997; 97US-0049914P.
XX 23-OCT-1997; 97US-0063595P.
XX
XX (SUGE-) SUGEN INC.
XX
XX Plowman GD, Clary D, Jallal B, Peles E, Onrust S, Markby D;
PI Courtneidge SA, App H, Hui TH;

ABZ34698;
04-FEB-2003 (first entry)
Coding sequence SEQ ID 56, downregulated in osteogenesis.
Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
osteoporosis; bone disease; downregulator; human; enzyme; ss.
Homo sapiens.
WO200281745-A2.
17-OCT-2002.
05-APR-2002; 2002WO-1B002211.
05-APR-2001; 2001US-0281400P.
(AVET) AVENTIS PHARMA SA.
Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;
Connolly T, Jackson A, Bushnell SE, Rawadi G;
WPI; 2003-058567/05.
Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
useful for bone disease therapy in subject.
Claim 27; Page 107; 237pp; English.
The present invention relates to novel nucleotide sequences, which are
differentially expressed in models of osteogenesis upon being put in
contact with a stimulator of osteogenesis. The present sequence is one
such sequence. This sequence can be used for diagnosing osteoporosis/bone
disease in a patient, promoting osteogenesis and/or preventing
osteoporosis/bone disease. The present sequence encodes an intracellular
enzyme
Sequence 1761 BP; 375 A; 509 C; 510 G; 367 T; 0 U; 0 Other;
Query Match 1.6%; Score 19; DB 7; Length 1761;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 68 TCCTCGCTCGCTGCT 86
DB 1078 TCCTCGCTCGCTGCT 1096
RESULT 93
ABZ34698
ID ABZ34698 standard; cDNA; 1833 BP.
AC ABZ34698;
XX ABZ34698;
DT 21-OCT-2002 (first entry)
DE Human cDNA encoding a Na/Cl-dependent organic solute transporter.
XX Human, ss; gene; Na/Cl-dependent organic solute cotransporter;
KW chromosome 5; neuron tumour.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 1..1833
FT /*tag= a
FT /product= "Na/Cl-dependent organic solute cotransporter"
XX WO200253741-A2.
XX 11-JUL-2002.

18-APR-2002.
12-OCT-2001; 2001WO-US031922.
12-OCT-2000; 2000US-0240113P.
16-OCT-2000; 2000US-0240632P.
16-OCT-2000; 2000US-0240637P.
16-OCT-2000; 2000US-0240648P.
16-OCT-2000; 2000US-0240662P.
16-OCT-2000; 2000US-0240663P.
16-OCT-2000; 2000US-0240703P.
16-OCT-2000; 2000US-0240732P.
16-OCT-2000; 2000US-0241190P.
18-JAN-2001; 2001US-0262455P.
(CURA-) CURAGEN CORP.
(MILL) MILLET I.
Grosse WM, Alsobrook JP, Lepley DM, Burgess CE, Mishra V;
Kekuda R, Li L, Padigaru M, Shimkets RA, Zerhusen BD, Spytek KA;
Edinger S, Gerlach V, Macdougall J, Stone D, Gunther E, Ellerman K;
WPI; 2002-444172/47.
P-FSDB; ABE78819.
New NOVX polypeptides and polynucleotides, useful for treating or
preventing a NOVX-associated disorder or a pathological state in a
subject, particularly a human, e.g. cardiomyopathy, atherosclerosis,
cancer or diabetes.
Claim 9; Page 71; 227pp; English.
The present invention describes novel human proteins designated NOVX
(where X is 1, 2a, 2b, 2c, 2d, 3, 4, 5, 6a, 6b, 7, 8, or 9). NOV1 is a
tyrosine-protein kinase 6-like protein; NOV2a-d are Keratin 4-like
proteins; NOV3 is a collagen-like protein; NOV4 is a cystatin B-like
protein; NOV5 is a serotonin receptor-like protein; NOV6a and NOV6b are
cold inducible glycoprotein 30-like proteins; NOV7 is a matrilin-2-like
protein; NOV8 is a leukocyte surface antigen (CD53)-like protein; and
NOV9 is a tyrosine kinase-like protein. NOVX sequences have cytostatic,
antiarteriosclerotic, cardiovascular, antidiabetic, immunosuppressive and
neuroprotective activities, and can be used in gene therapy. The NOVX
sequences can be used in therapeutics, particularly for treating,
preventing or alleviating a NOVX-associated disorder or a pathological
state in a subject, particularly a human. These disorders include
cardiomyopathy, atherosclerosis, a disorder related to cell signal
processing and metabolic pathway modulation or diabetes. The NOVX
sequences are also useful for determining the presence of or
predisposition to a disease associated with altered levels of NOVX
polypeptide or nucleic acid, particularly cancer. The NOVX sequences are
especially useful in therapeutic or prophylactic applications for
neoplastic or neurological disorders, and in the treatment of
adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune
response, AIDS, asthma, Crohn's disease, multiple sclerosis or Graft
versus host disease. The present sequence encodes the human NOV9 protein
from the present invention. NOV9 is located to chromosome 20
Sequence 1580 BP; 277 A; 532 C; 450 G; 281 T; 0 U; 0 Other;
Query Match 1.6%; Score 19; DB 6; Length 1580;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 AGGGGCTTCCTCATCCGG 760
DB 455 AGGGGCTTCCTCATCCGG 473
RESULT 92
ABZ34698
ID ABZ34698 standard; cDNA; 1761 BP.
XX

XX 02-JAN-2002; 2002WO-US0000111.
 PF
 XX
 PR 03-JAN-2001; 2001US-00752821.
 PR 22-MAY-2001; 2001US-00861846.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Guegler K, Brandon RC, Di Francesco V, Beasley EM;
 XX
 DR WPI; 2002-583618/62.
 DR P-PSDB; ABG30878.
 XX
 DR
 XX
 PT Novel isolated human peptide useful for treating disorders characterized
 PT by the absence of, inappropriate or unwanted expression of the receptor
 PT protein, and as immunogens to raise antibodies.
 XX
 PS Claim 4; Fig 1; 89pp; English.
 XX
 CC The invention relates to an isolated human peptide (S1) comprising an
 CC Na/Cl-dependent organic solute cotransporter family member, an allelic
 CC variant or orthologue of (S1) encoded by a polynucleotide hybridising
 CC under stringent conditions to the opposite strand of a sequence appearing
 CC as ABK89238 and ABK89239, or a fragment of (S1) with at least 10
 CC contiguous amino acids. A pharmaceutical composition comprising a
 CC modulator of S1 is useful for treating a disease or condition mediated by
 CC a human protease. S1 and its encoding nucleic acid are useful as models
 CC for the development of human therapeutics, for identifying therapeutic
 CC proteins, as targets for development of human therapeutic agents. S1 is
 CC useful to raise antibodies or to elicit another immune response, as a
 CC reagent in assays designed to quantitatively determine levels of the
 CC protein in biological fluids, as markers for tissues in which the
 CC corresponding protein is preferentially expressed, in drug screening
 CC assays, in cell-based or cell-free systems, to identify compounds that
 CC modulate cotransporter activity of the protein in its natural state, or
 CC an altered form that causes the specific disease or pathology associated
 CC with the receptor (e.g. neuron tumour), as bait proteins in a two-hybrid
 CC or three-hybrid assay, to provide a target for diagnosing a disease or
 CC predisposition to disease mediated by the peptide, in pharmacogenomic
 CC analysis, and for treating a disorder characterised by an absence of, in
 CC appropriate or unwanted expression of S1. An anti-S1 antibody is useful
 CC for isolating and purifying S1, to detect S1 in cells or tissues, in
 CC situ, in vitro or in a cell lysate or supernatant, to assess abnormal
 CC tissue distribution or abnormal expression during development or
 CC progression of a biological condition, and for inhibiting a protein
 CC function. S1 nucleic acids are useful as probes, primers, and chemical
 CC intermediates, in biological assays, for constructing recombinant
 CC vectors, host cells and transgenic animals, for expressing antigenic
 CC portions of the S1 proteins, or for designing ribozymes. A vector
 CC comprising the S1 nucleic acid is useful for producing an S1 protein or
 CC peptide, for conducting cell-based assays involving the protein or its
 CC fragment, for identifying protein mutants whose functions are affected,
 CC and to produce non-human transgenic animals. The gene for the
 CC cotransporter is located on chromosome 5. The present sequence encodes
 CC the cotransporter protein

XX SQ Sequence 1833 BP; 321 A; 530 G; 403 T; 0 U; 0 Other;
 Query Match 1.6%; Score 19; DB 6; Length 1833;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTCATCC 758
 DB 157 GGAGGGGCTTCTCATCC 175
 |||||
 |||||

RESULT 94
 ADA53009
 ID ADA53009 standard; cDNA; 2120 BP.
 XX
 AC ADA53009;

XX 20-NOV-2003 (first entry)
 XX Human coding sequence, SEQ ID 577.
 XX
 KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX EPI293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 XX 21-MAR-2002; 2002EP-00006586.
 PF
 XX 14-SEP-2001; 2001JP-00328381.
 PR
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 PA
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tameshika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 DR WPI; 2003-395539/38.
 DR P-PSDB; ADA54648.
 XX
 CC New polynucleotides encoding full-length polypeptides, e.g. secretory
 CC and/or membrane proteins, useful for developing medicines for diseases in
 CC PT which the gene is involved, or as target molecules for gene therapy.
 CC
 XX Claim 1; SEQ ID NO 577; 205pp; English.
 PS
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SQ Sequence 2120 BP; 370 A; 668 C; 634 G; 448 T; 0 U; 0 Other;
 Query Match 1.6%; Score 19; DB 7; Length 2120;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTCATCC 758
 DB 280 GGAGGGGCTTCTCATCC 298
 |||||
 |||||

RESULT 95
 ADE28356
 ID ADE28356 standard; cDNA; 2130 BP.
 XX
 AC ADE28356;
 XX
 XX 29-JAN-2004 (first entry)
 XX Human KPP cDNA - SEQ ID 67.
 DE
 XX
 XX kinase; phosphatase; KPP; hepatotropic; antiarteriosclerotic;
 KW antipsoriatic; cytostatic; haemostatic; muscular; cerebroprotective;
 KW nootropic; ophthalmological; anticonvulsant; vasotropic; neuroprotective;
 KW antiparkinsonian; antiasthmatic; antianaemic; antiarthritic;
 KW antidiabetic; antiinflammatory; osteopathic; antiasthmatic;
 KW antirheumatic; dermatological; virucide; antibacterial; fungicide;
 KW antiparasitic; protozoacide; antineoplastic; antitumor; cardiovascular;
 KW antiarteriosclerotic; immunosuppressive; cell proliferative; cirrhosis;
 KW hepatitis; arteriosclerosis; psoriasis; primary thrombocytopenia; cancer;
 KW developmental; renal tubular acidosis; Becker's muscular dystrophy;

Pick's disease; cataract; epilepsy; ischaemic cerebrovascular; stroke; Alzheimer's; Parkinson's; dementia; autoimmune; inflammatory; AIDS; allergy; anaemia; asthma; diabetes mellitus; bronchitis; osteoporosis; osteoarthritis; rheumatoid arthritis; contact dermatitis; gout; lipid disorder; cholestasis; Gaucher's; diabetes; atherosclerosis; liver; viral; bacterial; fungal; parasitic; protozoan; helminthic infection; trauma; gene therapy; human; ss; gene.

XX OS Homo sapiens.
XX XX WO2003080805-A2.
XX PD 02-OCT-2003.
XX XX 18-MAR-2003; 2003WO-US008715.
XX PR 19-MAR-2002; 2002US-0366088P.
XX PR 29-MAR-2002; 2002US-0369248P.
XX PA (INCY-) INCYTE CORP.
XX PI Chien D, Jin P, Hawkins PR, Baughn MR, Becha SD, Chang H;
PI Ding L, Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ, Griffin JA;
PI Gururajan R, Hafalia AJA, Ison CH, Kable AE, Khare R, Lee SY;
PI Lee EA, Lu Y, Marquis JP, Lehr-Wason PM, Ramkumar J, Richardson TW;
PI Swarnakar A, Tran UK, Chawla NK, Yao MG, Yue H, Bhatia U;
PI Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI; 2004-011523/01.
XX P-PSDB; ADE28304.
XX New human kinases and phosphatases, and polynucleotides encoding them,
XX PT useful for treating, preventing or diagnosing e.g. cell proliferative
XX disorders, inflammatory, autoimmune, viral, bacterial, parasitic or
XX fungal diseases.
XX Claim 5; SEQ ID NO 67; 340pp; English.

XX The invention relates to a novel isolated kinase and phosphatase (KPP)
XX polypeptide. The polypeptide of the invention demonstrates hepatotropic,
XX antiarteriosclerotic, antipsoriatic, cytostatic, haemostatic, muscular,
XX cerebroprotective, neurotropic, ophthalmological, anticonvulsant,
XX vasotropic, neuroprotective, antiparkinsonian, antiasthmatic,
XX anianemic, antiasthmatic, antidiabetic, antiinflammatory, osteopathic,
XX antiarthritic, antineumatic, dermatological, virucide, antibacterial,
XX fungicide, antiparasitic, protozoacide, antihelminthic, antigout,
XX cardiovascular, antiarteriosclerotic and immunosuppressive activities.
XX The KPP polypeptides may be useful for diagnosing, treating or preventing
XX cell proliferative disorders including cirrhosis, hepatitis,
XX arteriosclerosis, psoriasis, primary thrombocytopenia and cancer,
XX developmental disorders such as renal tubular acidosis, Becker's muscular
XX dystrophy, gonadal dysgenesis, hypothyroidism or seizures, neurological
XX disorders e.g. Pick's disease, cataract, epilepsy, ischaemic
XX cerebrovascular disease, stroke, Alzheimer's disease, Parkinson's disease
XX or dementia, autoimmune or inflammatory disorders including AIDS,
XX allergies, anaemia, asthma, diabetes mellitus, bronchitis, osteoporosis,
XX osteoarthritis, rheumatoid arthritis, contact dermatitis or gout and
XX lipid disorders such as cholestasis, Gaucher's disease, diabetes,
XX atherosclerosis or liver disease, as well as viral, bacterial, fungal,
XX parasitic, protozoan or helminthic infections and trauma. Furthermore,
XX the polypeptide may be utilised during gene therapy procedures. The
XX current sequence is that of the human KPP cDNA of the invention.

SQ Sequence 2130 BP; 394 A; 682 C; 671 G; 383 T; 0 U; 0 Other;
Query Match 1.6%; Score 19; DB 10; Length 2130;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 595 CTGAGGTGGAGACTGGTG 603

Db 1285 CTGAGGTGGAGACTGGTG 1303

RESULT 96

ACC46205

ID ACC46205 standard; cDNA; 2527 BP.

XX AC ACC46205;

XX DT 02-JUN-2003 (first entry)

XX DE Human dithp intracellular signalling protein-encoding cDNA.

XX KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;

XX KW cancer; cell proliferative disorder; autoimmune disorder;

XX KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;

XX KW neurological disorder; gastrointestinal disorder; transport disorder;

XX KW connective tissue disorder; drug screening; proteome analysis;

XX KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;

XX KW disease model; toxicological testing; transcript imaging;

XX KW intracellular signalling; gene; ss.

XX OS Homo sapiens.

XX PN WO200297031-A2.

XX PD 05-DEC-2002.

XX PF 27-MAR-2002; 2002WO-US010056.

XX PR 28-MAR-2001; 2001US-0279619P.

XX PR 29-MAR-2001; 2001US-0280067P.

XX PR 29-MAR-2001; 2001US-0280068P.

XX PR 16-MAY-2001; 2001US-0291280P.

XX PR 17-MAY-2001; 2001US-0291829P.

XX PR 19-MAY-2001; 2001US-0291849P.

XX PR 19-JUN-2001; 2001US-0299428P.

XX PR 20-JUN-2001; 2001US-0299776P.

XX PR 20-JUN-2001; 2001US-0300001P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

XX PI DuFour GB, Hillman JU, Yu JY, Tuason O, Yap PE, Anshey SR;

XX PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;

XX PI Paralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX DR WPI; 2003-129518/12.

XX DR P-PSDB; ABR41263.

XX PT Novel human diagnostic and therapeutic polypeptide useful for identifying

XX PT test compound which specifically binds to a polypeptide encoded by human

XX PT diagnostic and therapeutic polynucleotide, and to induce antibodies.

XX PS Claim 2; SEQ ID NO 126; 591pp; English.

XX CC The invention relates to novel human diagnostic and therapeutic

XX CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded

XX CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to

XX CC polynucleotide sequences at least 90% identical to the dithp cDNA

XX CC sequences of the invention; recombinant vectors, host cells and

XX CC transgenic organisms comprising a dithp nucleic acid sequence; the

XX CC recombinant organisms comprising dithp proteins; antibodies specific for DITHP

XX CC proteins; microarrays comprising dithp nucleic acid sequences; methods of

XX CC detecting dithp nucleotide and protein sequences; methods of screening

XX CC for compounds which specifically bind a DITHP protein; and methods of

XX CC assessing the toxicity of test compounds using a dithp hybridisation

XX CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the

XX CC diagnosis of a wide variety of conditions including cancer and other cell

XX CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,

XX CC viral, fungal or parasitic infections; hormonal disorders; metabolic

XX CC disorders; neurological disorders; gastrointestinal disorders; transport

XX CC disorders; and connective tissue disorders. They may also be used to

XX CC screen for modulators of protein activity or gene expression. DITHP

CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock in
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a dithp cDNA encoding a DITHP protein
 CC which has intracellular signalling activity. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2527 BP; 697 A; 568 C; 593 G; 569 T; 0 U; 0 Other;
 Query Match 1.6%; Score 19; DB 7; Length 2527;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 751 CCTCATCCGGAGGCCAG 769
 DB 555 CCTCATCCGGAGGCCAG 573

RESULT 97
 ACA56503
 ID ACA56503 standard; cDNA; 2771 BP.
 XX AC ACA56503;

DT 06-JUN-2003 (first entry)

XX Human signalling pathway polynucleotide probe SEQ ID NO 1101.

DE Human; probe; ss; array element; Parkinson's disease;
 KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
 KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.

XX Homo sapiens.

OS US6500938-B1.

PN 31-DEC-2002.

PD 30-JAN-1998; 98US-00016434.

PF 30-JAN-1998; 98US-00016434.

PR (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Seilhamer JJ;

XX WPI; 2003-352189/33.

XX Combination of polynucleotide probes, useful as array elements in a
 PT microarray for monitoring the expression of a number of target
 PT polynucleotides.

PS Claim 1; SEQ ID NO 1101; 65pp; English.

XX The invention relates to a combination which, comprises a number of
 CC polynucleotide probes comprising a sequence selected from one of the 1490
 CC sequences mentioned in the specification. The combination is useful as an
 CC array element in a microarray for monitoring the expression of a number
 CC of target polynucleotides. The microarray is particularly useful in the
 CC diagnosis and treatment of cancer and immunopathology and neuropathology.
 CC The microarray is useful in diagnostics and treatment regimens, drug
 CC discovery and development, toxicological and carcinogenicity studies,
 CC forensics and pharmacogenetics. The microarray is also useful for
 CC monitoring progression of diseases and for developing sophisticated
 CC profiles for the effects of currently available therapeutic drugs. The
 CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs

CC array can detect changes in expression in a large number of genes coding
 CC for different signaling pathway populations which can be used to diagnose
 CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
 CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
 CC and Parkinson's disease. The present sequence represents a polynucleotide
 CC probe of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=06500938B1

XX SQ Sequence 2771 BP; 534 A; 885 C; 814 G; 538 T; 0 U; 0 Other;
 Query Match 1.6%; Score 19; DB 7; Length 2771;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 585 CTCGAGATGGAGACTGGTG 603

DB 1305 CTCGAGATGGAGACTGGTG 1323

RESULT 98
 AAT60434
 ID AAT60434 standard; cDNA; 2774 BP.
 XX AC AAT60434;

DT 09-JUL-1997 (first entry)

XX Human intracellular tyrosine kinase Tnki-alpha cDNA.

DE Tyrosine kinase; Tnki-alpha; signal transduction; cell transformation;
 KW cell proliferation; haematopoietic cell; bone marrow; cancer;
 KW gene therapy; diagnosis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 117..2102
 FT /*tag= a

XX WO9713846-A1.

XX 17-APR-1997.

XX 11-OCT-1996; 96WO-US016359.

XX 12-OCT-1995; 95US-0005286P.

XX (UJJO) UNIV JOHNS HOPKINS.

XX Civin CI, Small D, Hoehn GT;

XX WPI; 1997-235892/21.

XX P-PSDB; AAW15565.

XX Tnki intracellular tyrosine kinase and its splice variant - useful in
 PT gene therapy to inhibit cell transformation, stimulate haematopoietic
 PT cells etc. and for diagnosis.

PS Claim 13; Page 46-47; 69pp; English.

XX A cDNA clone (AAT60434) codes for splice variant Tnki-alpha (AAW15565) of
 CC a novel human intracellular tyrosine kinase (see also AAT60433)
 CC designated Tnki. Its sequence is identical to that of Tnki except for the
 CC absence of nucleotides 1359-1374 of Tnki cDNA. Both cDNA clones were
 CC obtd. from human umbilical cord blood haematopoietic/ progenitor cells by
 CC PCR amplification using degenerate primers (see also AAT60434-35) that
 CC target conserved regions of tyrosine kinases, and use of the PCR product
 CC to design Tnki-specific primers (see also AAT60437-43) that were used for
 CC 3- and 5' RACE. The Tnki gene was mapped to chromosome 17p13.1, near the
 CC p33 locus. Tnki or Tnki-alpha produced by in vivo expression (Gene

CC GTPase activity of a p21 ras-like protein and to regulate foetal and post
 CC - natal development. Tnki-alpha nucleic acids can also be used for
 CC expression of recombinant Tnki-alpha and as probes/primers for isolation
 CC of related genes or for analysis of gene mutations, such as for diagnosis
 CC of cancer or susceptibility to it
 XX
 SQ Sequence 2774 BP; 550 A; 878 C; 813 G; 533 T; 0 U; 0 Other;

Query Match 1.6%; Score 19; DB 2; Length 2774;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 585 CTGAGGATGGAGACTGGTG 603
 DB 1305 CTGAGGATGGAGACTGGTG 1323

RESULT 99
 AAT60433
 ID AAT60433 standard; cDNA; 2789 BP.

XX AC AAT60433;
 DT 09-JUL-1997 (first entry)
 XX DE Human intracellular tyrosine kinase Tnki cDNA.

XX Tyrosine kinase; Tnki; signal transduction; cell transformation;
 KW cell proliferation; haematopoietic cell; bone marrow; cancer;
 KW gene therapy; diagnosis; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT CDS 117..2117
 FT /*tag= a

XX PN WO9713846-A1.

XX PD 17-APR-1997.

XX PF 11-OCT-1996; 96WO-US016359.

XX PR 12-OCT-1995; 95US-0005286P.

XX PA (UWJO) UNIV JOHNS HOPKINS.

XX PI Civin CI, Small D, Hoehn GT;

XX WPI; 1997-235882/21.

XX P-PSDB; AAW15564.

XX Tnki intracellular tyrosine kinase and its splice variant - useful in
 PT gene therapy to inhibit cell transformation, stimulate haematopoietic
 PT cells etc. and for diagnosis.

XX Claim 13; Fig 1; 69pp; English.

XX A cDNA clone (AAT60433) codes for a novel human intracellular tyrosine
 CC kinase (AAW15564) designated Tnki. cDNA derived from human umbilical cord
 CC blood haematopoietic/progenitor cells was subjected to PCR amplification
 CC using degenerate primers (see also AAT60434-35) that target conserved
 CC regions of tyrosine kinases. A 250 bp product was sequenced and Tnki-
 CC specific primers (see also AAT60437-43) were used for 3' and 5'RACE to
 CC isolate the full sequence sequence for Tnki and for its splice variant
 CC (see also AAT60434). The 2790 cDNA has been deposited in pBluescript KS(-
 CC) as ATCC 69924 (in E. coli DH5alpha). The Tnki gene was mapped to
 CC chromosome 17p13.1, near the p53 locus. Tnki produced by in vivo
 CC expression (gene therapy) can be used to inhibit cell transformation by
 CC regulating the GTPase activity of a p21 ras-like protein and to regulate
 CC foetal and post-natal development. Tnki nucleic acids can also be used
 CC for expression of recombinant Tnki and as probes/primers for isolation of
 CC related genes or for analysis of Tnki gene mutations, such as for

CC diagnosis of cancer or susceptibility to it

XX Sequence 2789 BP; 552 A; 885 C; 814 G; 538 T; 0 U; 0 Other;

SQ Query Match 1.6%; Score 19; DB 2; Length 2789;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 585 CTGAGGATGGAGACTGGTG 603

DB 1305 CTGAGGATGGAGACTGGTG 1323

RESULT 100

AAS80650

ID AAS80650 standard; cDNA; 3127 BP.

XX AC AAS80650;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #16454.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSB-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG16463.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1; SEQ ID NO 16454; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 3127 BP; 539 A; 986 C; 1041 G; 561 T; 0 U; 0 Other;
 Query Match 1.6%; Score 19; DB 5; Length 3127;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 AGGGGCTTCCTCATCCGG 760
 |||||
 Db 1099 AGGGGCTTCCTCATCCGG 1117

Search completed: March 25, 2004, 06:17:24
 Job time : 557 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 05:47:59 ; Search time 104 Seconds
(without alignments)
6312.566 Million cell updates/sec

Title: US-09-939-853A-74
Perfect score: 1183
Sequence: 1 agctagctccaggagcc.....tctcttgatgagcctag 1183

Scoring table: OLIGO_NUC

Searched: 682709 seqs, 277475446 residues
Gapop 60.0, Gapext 60.0

Word size : 0
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 300 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	1.7	539	4	US-09-621-976-10381
2	20	1.7	675	1	US-08-707-793A-3
3	20	1.7	675	1	US-08-707-792A-3
4	20	1.7	2129	4	US-09-016-434-1452
5	20	1.7	786431	4	US-09-751-389-3
6	19	1.6	1467	4	US-09-579-182-2
7	19	1.6	1548	4	US-09-099-053-1
8	19	1.6	2771	4	US-09-016-434-1101
9	18	1.5	1438	3	US-09-187-331-4
10	18	1.5	1438	3	US-09-470-946-4
11	18	1.5	1669	3	US-09-318-448-8
12	18	1.5	3090	3	US-09-276-531-78
13	18	1.5	70000	4	US-09-851-896-3
14	17	1.4	351	3	US-09-046-479-1
15	17	1.4	351	3	US-08-822-897C-1
16	17	1.4	351	4	US-09-608-810A-3
17	17	1.4	351	4	US-09-404-417A-1
18	17	1.4	435	4	US-09-252-991A-6817
19	17	1.4	439	4	US-09-222-575-172
20	17	1.4	439	4	US-09-389-681-172
21	17	1.4	439	4	US-09-620-405B-172
22	17	1.4	439	4	US-09-339-338-172
23	17	1.4	439	4	US-09-433-828B-172
24	17	1.4	439	4	US-09-604-287A-172
25	17	1.4	439	4	US-09-285-480-172
26	17	1.4	439	4	US-09-834-759-172
27	17	1.4	445	4	US-09-702-705-1598

4	445	4	US-09-736-457-1598	Sequence 1598, Ap
4	445	4	US-09-614-124B-1598	Sequence 1598, Ap
4	445	4	US-09-671-325-1598	Sequence 1598, Ap
4	541	4	US-09-220-132-10	Sequence 10, Appl
4	566	4	US-09-621-976-1574	Sequence 1574, Ap
4	627	4	US-09-328-352-1086	Sequence 1086, Ap
4	643	4	US-09-833-381-1262	Sequence 1262, Ap
4	651	4	US-09-016-434-1255	Sequence 1255, Ap
4	674	4	US-09-621-976-87	Sequence 87, Appl
4	759	4	US-09-489-039A-5493	Sequence 5493, Ap
4	894	4	US-09-434-354-1	Sequence 1, Appl
4	941	4	US-09-634-238-89	Sequence 89, Appl
4	1089	1	US-08-154-915-1	Sequence 1, Appl
2	1089	2	US-08-464-517-37	Sequence 37, Appl
2	1089	2	US-08-246-361A-37	Sequence 37, Appl
2	1089	2	US-08-463-772-37	Sequence 37, Appl
5	1089	5	PCT-US93-09945-1	Sequence 1, Appl
4	1747	4	US-09-566-921-66	Sequence 66, Appl
4	1748	4	US-08-765-889C-1	Sequence 1, Appl
4	1748	4	PCT-US95-07855-1	Sequence 1, Appl
4	1953	4	US-09-252-991A-6975	Sequence 6975, Ap
4	2118	4	US-09-252-991A-6727	Sequence 6727, Ap
2	2180	2	US-08-755-559-2	Sequence 2, Appl
3	2180	3	US-09-210-474-2	Sequence 2, Appl
4	2180	4	US-09-539-774-2	Sequence 2, Appl
4	2648	4	US-09-373-157-5	Sequence 5, Appl
4	2666	4	US-09-566-921-3	Sequence 3, Appl
4	2856	4	US-09-252-991A-6767	Sequence 6767, Ap
3	2874	3	US-09-179-558-54	Sequence 54, Appl
4	2874	4	US-09-722-825-54	Sequence 54, Appl
4	2874	4	US-09-722-467-54	Sequence 54, Appl
4	3059	4	US-09-722-708-54	Sequence 54, Appl
3	3059	3	US-09-179-558-61	Sequence 61, Appl
4	3059	4	US-09-722-825-61	Sequence 61, Appl
4	3059	4	US-09-722-487-61	Sequence 61, Appl
4	3059	4	US-09-722-708-61	Sequence 61, Appl
4	4139	4	US-09-620-312D-349	Sequence 349, App
3	8802	3	US-08-896-449A-1	Sequence 1, Appl
3	8802	3	US-09-132-652-1	Sequence 1, Appl
4	9840	4	US-09-534-638-1	Sequence 1, Appl
4	61663	4	US-09-453-702B-62	Sequence 62, Appl
3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
3	4411529	3	US-09-103-840A-1	Sequence 1, Appl
4	18	4	US-09-404-417A-8	Sequence 8, Appl
4	211	4	US-09-833-381-378	Sequence 378, App
4	251	4	US-09-602-877A-93	Sequence 93, Appl
3	265	3	US-09-071-710-9	Sequence 9, Appl
3	265	3	US-09-525-397-9	Sequence 9, Appl
4	278	4	US-09-621-976-16442	Sequence 16442, A
4	283	4	US-09-313-294A-3534	Sequence 3534, Ap
3	288	3	US-09-071-710-10	Sequence 10, Appl
3	288	3	US-09-525-397-10	Sequence 10, Appl
1	310	1	US-08-235-838-9	Sequence 9, Appl
2	310	2	US-08-465-473B-9	Sequence 9, Appl
2	321	2	US-08-888-366-21	Sequence 21, Appl
4	352	4	US-09-641-638-433	Sequence 433, Appl
4	352	4	US-09-641-638-434	Sequence 434, App
4	391	4	US-09-621-976-18404	Sequence 18404, A
4	399	4	US-09-489-039A-3397	Sequence 3397, Ap
2	408	2	US-09-564-329A-10	Sequence 10, Appl
2	423	2	US-08-822-028-62	Sequence 62, Appl
3	423	3	US-08-479-285-62	Sequence 62, Appl
3	423	3	US-09-503-653A-62	Sequence 360, App
3	439	3	US-09-042-353-360	Sequence 288, App
4	439	4	US-08-758-417A-208	Sequence 10604, A
4	480	4	US-09-621-976-10604	Sequence 2590, Ap
4	497	4	US-09-621-976-2590	Sequence 1532, Ap
4	546	4	US-09-252-991A-1532	Sequence 491, App
4	579	4	US-09-328-352-491	Sequence 11119, A
4	594	4	US-09-252-991A-11119	Sequence 12620, A
4	615	4	US-09-252-991A-12620	Sequence 9984, Ap
4	654	4	US-09-252-991A-9984	Sequence 1241, A

C 101	16	1.4	748	1	US-08-235-838-10	Sequence 10, Appl	174	16	1.4	2662	4	US-08-323-430-14	Sequence 14, Appl
C 102	16	1.4	748	2	US-08-465-473B-10	Sequence 10, Appl	175	16	1.4	2733	1	US-08-371-001-14	Sequence 14, Appl
C 103	16	1.4	754	4	US-09-833-381-968	Sequence 968, Appl	176	16	1.4	2733	5	PCT-US96-00331-14	Sequence 14, Appl
C 104	16	1.4	772	3	US-09-020-956-11	Sequence 11, Appl	177	16	1.4	2859	2	US-08-506-340A-2	Sequence 2, Appl
C 105	16	1.4	772	3	US-09-030-607-11	Sequence 11, Appl	178	16	1.4	2864	4	US-09-409-180A-2	Sequence 2, Appl
C 106	16	1.4	772	4	US-09-439-313-11	Sequence 11, Appl	179	16	1.4	2872	4	US-09-327-487A-2	Sequence 2, Appl
C 107	16	1.4	772	4	US-09-352-616A-11	Sequence 11, Appl	180	16	1.4	2904	4	US-09-685-166A-703	Sequence 703, Appl
C 108	16	1.4	772	4	US-09-232-149A-11	Sequence 11, Appl	181	16	1.4	2904	4	US-09-685-166A-703	Sequence 703, Appl
C 109	16	1.4	772	4	US-09-159-812-11	Sequence 11, Appl	182	16	1.4	3146	4	US-09-620-312D-277	Sequence 277, Appl
C 110	16	1.4	772	4	US-09-636-215-11	Sequence 11, Appl	183	16	1.4	3410	3	US-09-020-956-110	Sequence 110, Appl
C 111	16	1.4	772	4	US-09-685-166A-11	Sequence 11, Appl	184	16	1.4	3410	3	US-09-030-607-110	Sequence 110, Appl
C 112	16	1.4	772	4	US-09-115-453-11	Sequence 11, Appl	185	16	1.4	3410	4	US-09-439-313-110	Sequence 110, Appl
C 113	16	1.4	772	4	US-09-688-489-11	Sequence 11, Appl	186	16	1.4	3410	4	US-09-352-616A-110	Sequence 110, Appl
C 114	16	1.4	819	1	US-08-792-019B-4	Sequence 4, Appl	187	16	1.4	3410	4	US-09-602-877A-100	Sequence 100, Appl
C 115	16	1.4	819	3	US-08-988-819-4	Sequence 4, Appl	188	16	1.4	3410	4	US-09-232-149A-110	Sequence 110, Appl
C 116	16	1.4	819	3	US-09-016-534-4	Sequence 4, Appl	189	16	1.4	3410	4	US-09-159-812-110	Sequence 110, Appl
C 117	16	1.4	847	1	US-08-053-131-184	Sequence 184, Appl	190	16	1.4	3410	4	US-09-636-215-110	Sequence 110, Appl
C 118	16	1.4	847	1	US-08-096-762-184	Sequence 184, Appl	191	16	1.4	3410	4	US-09-685-166A-110	Sequence 110, Appl
C 119	16	1.4	847	3	US-09-042-353-47	Sequence 47, Appl	192	16	1.4	3410	4	US-09-115-453-110	Sequence 110, Appl
C 120	16	1.4	847	4	US-08-758-417A-312	Sequence 312, Appl	193	16	1.4	3410	4	US-09-688-489-110	Sequence 110, Appl
C 121	16	1.4	908	3	US-08-718-388-1	Sequence 1, Appl	194	16	1.4	3524	4	US-09-077-940A-3	Sequence 3, Appl
C 122	16	1.4	909	4	US-09-540-236-1829	Sequence 1829, Appl	195	16	1.4	3530	3	US-08-704-711A-10	Sequence 10, Appl
C 123	16	1.4	924	1	US-08-468-709B-1	Sequence 1, Appl	196	16	1.4	3530	4	US-09-521-220-10	Sequence 10, Appl
C 124	16	1.4	924	2	US-08-241-664B-1	Sequence 1, Appl	197	16	1.4	3831	4	US-09-360-394C-1	Sequence 1, Appl
C 125	16	1.4	924	4	US-09-640-173-174	Sequence 174, Appl	198	16	1.4	3839	3	US-09-056-105-14	Sequence 14, Appl
C 126	16	1.4	924	4	US-09-640-173-174	Sequence 174, Appl	199	16	1.4	3931	4	US-08-956-171B-342	Sequence 342, Appl
C 127	16	1.4	924	5	PCT-US93-03936-1	Sequence 1, Appl	200	16	1.4	4034	4	US-09-636-215-704	Sequence 704, Appl
C 128	16	1.4	927	3	US-09-147-915-2	Sequence 2, Appl	201	16	1.4	4034	4	US-09-685-166A-704	Sequence 704, Appl
C 129	16	1.4	1065	3	US-08-875-811-56	Sequence 56, Appl	202	16	1.4	4394	4	US-09-620-312D-237	Sequence 237, Appl
C 130	16	1.4	1125	4	US-09-218-489-1	Sequence 1, Appl	203	16	1.4	4832	4	US-08-851-567B-58	Sequence 58, Appl
C 131	16	1.4	1230	4	US-09-252-991A-3665	Sequence 3665, Appl	204	16	1.4	4834	4	US-09-636-215-702	Sequence 702, Appl
C 132	16	1.4	1336	3	US-08-718-388-2	Sequence 2, Appl	205	16	1.4	4894	4	US-09-685-166A-702	Sequence 702, Appl
C 133	16	1.4	1386	4	US-09-252-991A-1489	Sequence 1489, Appl	206	16	1.4	5330	4	US-09-023-905A-1	Sequence 1, Appl
C 134	16	1.4	1572	4	US-09-489-039A-5714	Sequence 5714, Appl	207	16	1.4	5430	3	US-09-012-515A-11	Sequence 11, Appl
C 135	16	1.4	1639	4	US-09-620-312D-317	Sequence 317, Appl	208	16	1.4	5430	3	US-08-360-144A-11	Sequence 11, Appl
C 136	16	1.4	1641	4	US-09-328-352-1458	Sequence 1458, Appl	209	16	1.4	5430	4	US-09-012-504A-11	Sequence 11, Appl
C 137	16	1.4	1692	4	US-09-540-236-978	Sequence 978, Appl	210	16	1.4	5430	4	US-09-012-399A-11	Sequence 11, Appl
C 138	16	1.4	1895	1	US-08-361-920-20	Sequence 20, Appl	211	16	1.4	5524	4	US-09-844-497-3	Sequence 3, Appl
C 139	16	1.4	1895	1	US-08-479-939-20	Sequence 20, Appl	212	16	1.4	6822	4	US-09-426-998-3	Sequence 3, Appl
C 140	16	1.4	1695	1	US-08-483-433-20	Sequence 20, Appl	213	16	1.4	6976	4	US-09-636-215-705	Sequence 705, Appl
C 141	16	1.4	1716	4	US-09-674-677-6	Sequence 6, Appl	214	16	1.4	6976	4	US-09-685-166A-705	Sequence 705, Appl
C 142	16	1.4	1717	1	US-08-468-709B-6	Sequence 6, Appl	215	16	1.4	7653	4	US-08-471-112A-1	Sequence 1, Appl
C 143	16	1.4	1717	2	US-08-241-664B-6	Sequence 6, Appl	216	16	1.4	7741	4	US-09-426-998-4	Sequence 4, Appl
C 144	16	1.4	1717	5	PCT-US93-03936-6	Sequence 6, Appl	217	16	1.4	7824	3	US-08-718-388-6	Sequence 6, Appl
C 145	16	1.4	1770	4	US-09-252-991A-1639	Sequence 1639, Appl	218	16	1.4	7824	5	PCT-US95-06722-11	Sequence 11, Appl
C 146	16	1.4	1824	4	US-09-016-434-1425	Sequence 1425, Appl	219	16	1.4	8285	4	US-09-732-023-3	Sequence 3, Appl
C 147	16	1.4	1825	4	US-09-023-655-1061	Sequence 1061, Appl	220	16	1.4	8598	4	US-08-305-790B-1	Sequence 1, Appl
C 148	16	1.4	1839	4	US-09-252-991A-1687	Sequence 1687, Appl	221	16	1.4	9046	1	US-08-227-536-1	Sequence 1, Appl
C 149	16	1.4	1868	4	US-09-739-455-1	Sequence 1, Appl	222	16	1.4	9046	5	PCT-US95-04682-1	Sequence 1, Appl
C 150	16	1.4	1878	4	US-09-739-455-1	Sequence 1, Appl	223	16	1.4	11517	3	US-07-920-281C-1	Sequence 1, Appl
C 151	16	1.4	1882	4	US-09-620-312D-427	Sequence 427, Appl	224	16	1.4	11517	3	US-08-466-277-1	Sequence 1, Appl
C 152	16	1.4	1953	4	US-09-252-991A-3804	Sequence 3804, Appl	225	16	1.4	11725	2	US-08-756-506-1	Sequence 1, Appl
C 153	16	1.4	1994	4	US-09-398-395A-41	Sequence 41, Appl	226	16	1.4	11725	4	US-09-328-925-50	Sequence 50, Appl
C 154	16	1.4	1994	4	US-09-887-586A-41	Sequence 41, Appl	227	16	1.4	11827	4	US-09-739-455-3	Sequence 3, Appl
C 155	16	1.4	1994	4	US-09-895-752-41	Sequence 41, Appl	228	16	1.4	16382	3	US-08-718-388-8	Sequence 8, Appl
C 156	16	1.4	1994	4	US-09-903-012B-41	Sequence 41, Appl	229	16	1.4	24707	4	US-09-740-027-3	Sequence 3, Appl
C 157	16	1.4	1994	4	US-09-900-797-41	Sequence 41, Appl	230	16	1.4	34001	4	US-09-596-002-18	Sequence 18, Appl
C 158	16	1.4	2012	1	US-08-235-838-15	Sequence 15, Appl	231	16	1.4	51259	3	US-08-781-891-209	Sequence 209, Appl
C 159	16	1.4	2143	2	US-08-465-473B-15	Sequence 15, Appl	232	16	1.4	51259	4	US-09-618-166-209	Sequence 209, Appl
C 160	16	1.4	2143	3	US-09-071-710-15	Sequence 15, Appl	233	16	1.4	70000	4	US-09-851-896-3	Sequence 3, Appl
C 161	16	1.4	2143	3	US-09-525-397-15	Sequence 15, Appl	234	16	1.4	80246	3	US-09-078-294-4	Sequence 4, Appl
C 162	16	1.4	2152	3	US-09-071-710-16	Sequence 16, Appl	235	16	1.4	80595	3	US-09-078-294-3	Sequence 3, Appl
C 163	16	1.4	2152	3	US-09-253-397-16	Sequence 16, Appl	236	16	1.4	128779	4	US-09-497-855A-38	Sequence 38, Appl
C 164	16	1.4	2355	4	US-09-253-991A-3845	Sequence 3845, Appl	237	16	1.4	269223	4	US-09-596-002-41	Sequence 41, Appl
C 165	16	1.4	2360	4	US-09-023-655-1288	Sequence 1288, Appl	238	16	1.4	1830121	4	US-09-557-884-1	Sequence 1, Appl
C 166	16	1.4	2416	4	US-09-016-434-1264	Sequence 1264, Appl	239	16	1.4	1830121	4	US-09-643-950A-1	Sequence 30, Appl
C 167	16	1.4	2469	1	US-07-997-133-2	Sequence 2, Appl	240	15	1.3	27	3	US-09-253-396A-30	Sequence 39, Appl
C 168	16	1.4	2469	1	US-08-459-296-1	Sequence 1, Appl	241	15	1.3	28	1	US-08-467-420A-39	Sequence 39, Appl
C 169	16	1.4	2469	5	US-07-997-133-2	Sequence 2, Appl	242	15	1.3	28	1	US-08-470-110A-39	Sequence 39, Appl
C 170	16	1.4	2491	4	US-09-023-655-655	Sequence 655, Appl	243	15	1.3	28	2	US-08-667-769A-39	Sequence 39, Appl
C 171	16	1.4	2577	4	US-09-266-464-1	Sequence 1, Appl	244	15	1.3	28	2	US-08-940-371-39	Sequence 39, Appl
C 172	16	1.4	2577	4	US-09-016-434-1095	Sequence 1095, Appl	245	15	1.3	28	3	US-08-637-647-39	Sequence 39, Appl
C 173	16	1.4	2662	2	US-08-451-822A-14	Sequence 14, Appl	246	15	1.3	28	5	PCT-US95-17082A-39	Sequence 39, Appl


```

; GENERAL INFORMATION:
; APPLICANT: MASCO, ALICE
; APPLICANT: SALOWE, SCOTT P.
; APPLICANT: WISNIEWSKI, DOUGLAS
; TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,792A
; FILING DATE: 04-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Canata, Valerie J
; REGISTRATION NUMBER: 35,090
; REFERENCE/DOCKET NUMBER: 19524
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-3902
; TELEFAX: 908-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 675 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-707-792A-3

Query Match 1.7%; Score 20; DB 1; Length 675;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CTTCTCTCATCGGAGAGC 766
Db 443 CTTCTCTCATCGGAGAGC 462

RESULT 4
US-09-016-434-1452
; Sequence 1452, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1452:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g775207
; US-09-016-434-1452

Query Match 1.7%; Score 20; DB 4; Length 2129;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CTTCTCTCATCGGAGAGC 766
Db 508 CTTCTCTCATCGGAGAGC 527

RESULT 5
US-09-751-389-3
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GURGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
; US-09-751-389-3

Query Match 1.7%; Score 20; DB 4; Length 786431;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 693 TGAGCAGCGGAGAAAGCAGAG 712
Db 412751 TGAGCAGCGGAGAAAGCAGAG 412770

RESULT 6
US-09-579-182-2
; Sequence 2, Application US/09579182
; Patent No. 6500628

```

GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MM-161
; CURRENT APPLICATION NUMBER: US/09/579,182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-579-182-2

Query Match 1.6%; Score 19; DB 4; Length 1467;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 AGGGGCTTCCTCATCCGG 760
Db 423 AGGGGCTTCCTCATCCGG 441

RESULT 7
US-09-099-053-1
; Sequence 1, Application US/09099053
; Patent No. 6388063
; GENERAL INFORMATION:
; APPLICANT: Greg Plowman
; APPLICANT: Susan Oarust
; APPLICANT: David Markby
; APPLICANT: Sara Courtneidge
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: SAD RELATED DISORDERS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,053
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/049,914
; FILING DATE: June 18, 1997

ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 235/121
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1548 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-099-053-1

Query Match 1.6%; Score 19; DB 4; Length 1548;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 AGGGGCTTCCTCATCCGG 760
Db 471 AGGGGCTTCCTCATCCGG 489

RESULT 8
US-09-016-434-1101
; Sequence 1101, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4186
; INFORMATION FOR SEQ ID NO: 1101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENE BANK
; CLONE: g1256002
US-09-016-434-1101

Query Match 1.6%; Score 19; DB 4; Length 2771;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 585 CTGAGGATGGAGACTGGTG 603
Db 1305 CTGAGGATGGAGACTGGTG 1323

RESULT 9
US-09-187-331-4
; Sequence 4, Application US/09187331
; Patent No. 6043056
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.

APPLICANT: Guegler, Karl J.
APPLICANT: Gorgone, Gina A.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
FILE REFERENCE: PF-0631 US
CURRENT APPLICATION NUMBER: US/09/187,331
CURRENT FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 1438
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 2705267
US-09-187-331-4

Query Match 1.5%; Score 18; DB 3; Length 1438;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
|||||
DB 855 TACATCAGCCTGAATGAC 872

RESULT 10
US-09-470-946-4
Sequence 4, Application US/09470946
Patent No. 6358923
GENERAL INFORMATION:
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Gorgone, Gina A.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
FILE REFERENCE: PF-0631 US
CURRENT APPLICATION NUMBER: US/09/470,946
CURRENT FILING DATE: 1999-12-22
EARLIER APPLICATION NUMBER: US 09/187,331
EARLIER FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 1438
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 2705267
US-09-470-946-4

Query Match 1.5%; Score 18; DB 4; Length 1438;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
|||||
DB 855 TACATCAGCCTGAATGAC 872

RESULT 11
US-09-318-448-8
Sequence 8, Application US/09318448
Patent No. 6210950
GENERAL INFORMATION:
APPLICANT: Johnson, William G.
APPLICANT: Steenkos, Edward S.
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
FILE REFERENCE: 601-1-057
CURRENT APPLICATION NUMBER: US/09/318,448
CURRENT FILING DATE: 1999-05-25

NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1669
TYPE: DNA
ORGANISM: Homo sapiens
US-09-318-448-8

Query Match 1.5%; Score 18; DB 3; Length 1669;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 AGAAGAAATCTCTGCCA 433
|||||
DB 494 AGAAGAAATCTCTGCCA 511

RESULT 12
US-09-276-531-78/c
Sequence 78, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 3090 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT14
CLONE: 1595762
US-09-276-531-78

Query Match 1.5%; Score 18; DB 3; Length 3090;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1093 GGAGGAGTCTCTTCTCAG 1110
|||||
Db 398 GGAGGAGTCTCTTCTCAG 381

RESULT 13
US-09-851-896-3
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTIGENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)
; TITLE OF INVENTION: EXPRESSION
; FILE REFERENCE: RTS-0220
; CURRENT APPLICATION NUMBER: US/09/851,896
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; US-09-851-896-3

Query Match 1.5%; Score 18; DB 4; Length 70000;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 455 GTCCAGGCCAGGACCT 472
|||||
Db 60708 GTCCAGGCCAGGACCT 60725

RESULT 14
US-09-046-479-1/c
; Sequence 1, Application US/09046479
; Patent No. 6291653
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Deisher, Theresa A.
; TITLE OF INVENTION: MOTILIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,479
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A.
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 97-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...351

SEQUENCE CHARACTERISTICS:
LENGTH: 351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...351
OTHER INFORMATION:
NAME/KEY: sig_peptide
LOCATION: 1...69
OTHER INFORMATION:
NAME/KEY: mat_peptide
LOCATION: 70...351
OTHER INFORMATION:
US-09-046-479-1

Query Match 1.4%; Score 17; DB 3; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 GTCCAGCCAGCATGC 312
|||||
Db 57 GTCCAGCCAGCATGC 41

RESULT 15
US-08-822-897C-1/c
; Sequence 1, Application US/08822897C
; Patent No. 6360158
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Deisher, Theresa A.
; TITLE OF INVENTION: MOTILIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,897C
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A.
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 97-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...351

```
;
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: 1...69
; OTHER INFORMATION:
; NAME/KEY: mat_peptide
; LOCATION: 70...351
; OTHER INFORMATION:
;
US-08-822-897C-1
Query Match 1.4%; Score 17; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 GTCCAGCCAGCAGCATGC 312
Db 57 GTCCAGCCAGCAGCATGC 41

RESULT 16
US-09-608-810A-3/c
; Sequence 3, Application US/09608810A
; Patent No. 6420521
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: SCIP PEPTIDES
; FILE REFERENCE: 99-51
; CURRENT APPLICATION NUMBER: US/09/608,810A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/141,592
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(351)
; NAME/KEY: sig_peptide
; LOCATION: (1)...(69)
; NAME/KEY: mat_peptide
; LOCATION: (70)...(351)
US-09-608-810A-3

Query Match 1.4%; Score 17; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 GTCCAGCCAGCAGCATGC 312
Db 57 GTCCAGCCAGCAGCATGC 41

RESULT 17
US-09-404-417A-1/c
; Sequence 1, Application US/09404417A
; Patent No. 6627729
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Jaspers, Stephen R.
; TITLE OF INVENTION: TML PEPTIDES
; FILE REFERENCE: 97-04C1
; CURRENT APPLICATION NUMBER: US/09/404,417A
; CURRENT FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 351

;
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(351)
; NAME/KEY: sig_peptide
; LOCATION: (1)...(69)
; NAME/KEY: mat_peptide
; LOCATION: (70)...(351)
US-09-404-417A-1

Query Match 1.4%; Score 17; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 GTCCAGCCAGCAGCATGC 312
Db 57 GTCCAGCCAGCAGCATGC 41

RESULT 18
US-09-252-991A-6817/c
; Sequence 6817, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6817
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6817

Query Match 1.4%; Score 17; DB 4; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 CTGAGCAGGGAGAAAGC 708
Db 427 CTGAGCAGGGAGAAAGC 411

RESULT 19
US-09-222-575-172
; Sequence 172, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; FILE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/09/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (19)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (375)
```


; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (388)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (390)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (395)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (409)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (426)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (434)
; OTHER INFORMATION: Where n is a, c, g or t
US-09-222-575-172

Query Match 1.4%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 750 TCCTCATCCGGGAGAGC 766
Db 80 TCCTCATCCGGGAGAGC 96

RESULT 20
US-09-389-681-172
; Sequence 172, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-389-681-172

Query Match 1.4%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 750 TCCTCATCCGGGAGAGC 766
Db 80 TCCTCATCCGGGAGAGC 96

RESULT 21
US-09-620-405B-172
; Sequence 172, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Repler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-620-405B-172

Query Match 1.4%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 750 TCCTCATCCGGGAGAGC 766
Db 80 TCCTCATCCGGGAGAGC 96

RESULT 22
US-09-339-338-172
; Sequence 172, Application US/09339338A
; Patent No. 6573368
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-339-338-172

Query Match 1.4%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 750 TCCTCATCCGGGAGAGC 766
Db 80 TCCTCATCCGGGAGAGC 96

RESULT 23
US-09-433-826B-172
; Sequence 172, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-433-826B-172

Query Match 1.4%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGGAGGC 766
|||
Db 80 TCCTCATCCGGGAGGC 96

RESULT 24

US-09-604-287A-172
; Sequence 172, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-604-287A-172

Query Match 1.4%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGGAGGC 766
|||
Db 80 TCCTCATCCGGGAGGC 96

RESULT 25

US-09-285-480-172
; Sequence 172, Application US/09285480
; Patent No. 6590076
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C1
; CURRENT APPLICATION NUMBER: US/09/285,480
; CURRENT FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-285-480-172

Query Match 1.4%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGGAGGC 766
|||
Db 80 TCCTCATCCGGGAGGC 96

RESULT 26

US-09-834-759-172
; Sequence 172, Application US/09834759
; Patent No. 8680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-759-172

Query Match 1.4%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGGAGGC 766
|||
Db 80 TCCTCATCCGGGAGGC 96

RESULT 27

US-09-702-705-1598
; Sequence 1598, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary

```
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1598
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-1598

Query Match      1.4%; Score 17; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGGAGGC 766
   |||||
Db 59 TCCTCATCCGGGAGGC 75

RESULT 28
US-09-736-457-1598
; Sequence 1598, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1598
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1598

Query Match      1.4%; Score 17; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGGAGGC 766
   |||||
Db 59 TCCTCATCCGGGAGGC 75

RESULT 29
US-09-614-124B-1598
; Sequence 1598, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
```

```
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1598
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-614-124B-1598

Query Match      1.4%; Score 17; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGGAGGC 766
   |||||
Db 59 TCCTCATCCGGGAGGC 75

RESULT 30
US-09-671-325-1598
; Sequence 1598, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1598
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-1598

Query Match      1.4%; Score 17; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGGAGGC 766
   |||||
Db 59 TCCTCATCCGGGAGGC 75

RESULT 31
US-09-220-132-10
; Sequence 10, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
```

; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; CURRENT APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(541)
; OTHER INFORMATION: n = A,T,C or G
US-09-220-132-10

Query Match 1.4%; Score 17; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCTCATCGGGAGGC 766
Db 395 TCTCATCGGGAGGC 411

RESULT 32
US-09-621-976-1574
; Sequence 1574, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1574
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 176...439
; NAME/KEY: sig_peptide
; LOCATION: 176...247
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 6.09999990463257
; NAME/KEY: misc_feature
; LOCATION: 525
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-1574

Query Match 1.4%; Score 17; DB 4; Length 566;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1072 GTTTTCTGAAGTGCA 1088
Db 505 GTTTTCTGAAGTGCA 521

RESULT 33
US-09-328-352-1086/c
; Sequence 1086, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Berton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1086
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1086

Query Match 1.4%; Score 17; DB 4; Length 627;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1075 TTCTGAAGCTGCCACAG 1091
Db 117 TTCTGAAGCTGCCACAG 101

RESULT 34
US-09-833-381-1262
; Sequence 1262, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1262
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(643)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1262

Query Match 1.4%; Score 17; DB 4; Length 643;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCTGAATGA 1155
Db 156 TACATCAGCTGAATGA 172

RESULT 35
US-09-016-434-1255/c
; Sequence 1255, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERWITH
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1255:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g190878
US-09-016-434-1255

Query Match 1.4%; Score 17; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGAGAGC 766
DB 638 TCCTCATCCGGAGAGC 622

RESULT 36
US-09-621-976-87
Sequence 87, Application US/09621976
Patent No. 6639663
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 87
LENGTH: 674
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 221..673
NAME/KEY: sig_peptide
LOCATION: 221..268
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 7.30000019073486
OTHER INFORMATION: seq FULLTCLFITGTS/V5
US-09-621-976-87

Query Match 1.4%; Score 17; DB 4; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGA 1155
DB 302 TACATCAGCCTGAATGA 318

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERWITH
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1255:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g190878
US-09-016-434-1255

Query Match 1.4%; Score 17; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCCTCATCCGGAGAGC 766
DB 638 TCCTCATCCGGAGAGC 622

RESULT 36
US-09-621-976-87
Sequence 87, Application US/09621976
Patent No. 6639663
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 87
LENGTH: 674
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 221..673
NAME/KEY: sig_peptide
LOCATION: 221..268
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 7.30000019073486
OTHER INFORMATION: seq FULLTCLFITGTS/V5
US-09-621-976-87

Query Match 1.4%; Score 17; DB 4; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGA 1155
DB 302 TACATCAGCCTGAATGA 318

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 5493
LENGTH: 759
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5493

Query Match 1.4%; Score 17; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 542 GCCGAGCTGTCGCTGAG 558
DB 676 GCCGAGCTGTCGCTGAG 692

RESULT 37
US-09-489-039A-5493
Sequence 1, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 5493
LENGTH: 759
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5493

Query Match 1.4%; Score 17; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 542 GCCGAGCTGTCGCTGAG 558
DB 676 GCCGAGCTGTCGCTGAG 692

RESULT 38
US-09-434-354-1
Sequence 1, Application US/09434354
Patent No. 6562563
GENERAL INFORMATION:
APPLICANT: Murphy, Anne N.
APPLICANT: Clevenger, William
APPLICANT: Wiley, Sandra Eileen
APPLICANT: Andreyev, Alexander Y.
APPLICANT: Frigeri, Luciano G.
APPLICANT: Velicelebi, Gonul
APPLICANT: Davis, Robert E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETERMINING
INTERACTIONS OF MITOCHONDRIAL COMPONENTS, AND FOR
IDENTIFYING AGENTS THAT ALTER SUCH INTERACTIONS
FILE REFERENCE: 660088.433
CURRENT APPLICATION NUMBER: US/09/434,354
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 54
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 894
TYPE: DNA
ORGANISM: Homo sapien
US-09-434-354-1

Query Match 1.4%; Score 17; DB 4; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGAATCCCTAAGGAGCA 122
DB 178 AGAATCCCTAAGGAGCA 194

RESULT 39
US-09-634-238-89/c
Sequence 89, Application US/09634238
Patent No. 6544772
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Bloksberg, Leonard, N.
APPLICANT: Lubbers, Mark W.
```

APPLICANT: Dekker, James
APPLICANT: Christenson, Anna C.
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul W.
APPLICANT: Reid, Julian R.
APPLICANT: Coolbear, Timothy
TITLE OF INVENTION: Polynucleotides, materials incorporating
TITLE OF INVENTION: them and methods for using them.
FILE REFERENCE: 11000.1043U1
CURRENT APPLICATION NUMBER: US/09/634,238
CURRENT FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 422
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 89
LENGTH: 941
TYPE: DNA
ORGANISM: Lactobacillus rhamnosus
US-09-634-238-89

Query Match 1.4%; Score 17; DB 4; Length 941;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 888 TCACCTTCCCTCACTC 904
DB 916 TCACCTTCCCTCACTC 900

RESULT 40
US-08-154-915-1
Sequence 1, Application US/08154915
Patent No. 5618669
GENERAL INFORMATION:
APPLICANT: Beach, David
APPLICANT: Xiong, Yue
TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses
TITLE OF INVENTION: Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/154,915
FILING DATE: 19-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,997
FILING DATE: 17-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888
US-08-154-915-1

Query Match 1.4%; Score 17; DB 1; Length 1089;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 949 CTGCCTACTCAAGGAGC 965
DB 168 CTGCCTACTCAAGGAGC 184

RESULT 41
US-08-464-517-37
Sequence 37, Application US/08464517
Patent No. 5869640
GENERAL INFORMATION:
APPLICANT: Beach, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,517
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888
US-08-464-517-37

Query Match 1.4%; Score 17; DB 2; Length 1089;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 949 CTGCCTACTCAAGGAGC 965
Db 168 CTGCCTACTCAAGGAGC 184

RESULT 42

US-08-246-361A-37
; Sequence 37, Application US/08246361A
; Patent No. 5998582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,361A
FILING DATE: 19-MAY-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992

PRIOR APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992

ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent

REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 13..888

US-08-246-361A-37

Query Match 1.4%; Score 17; DB 2; Length 1089;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 949 CTGCCTACTCAAGGAGC 965
Db 168 CTGCCTACTCAAGGAGC 184

RESULT 43

US-08-463-772-37
; Sequence 37, Application US/08463772
; Patent No. 6066501

GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,772
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992

PRIOR APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992

ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent

REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 13..888

US-08-463-772-37

Query Match 1.4%; Score 17; DB 3; Length 1089;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 949 CTGCCTACTCAAGGAGC 965
Db 168 CTGCCTACTCAAGGAGC 184

RESULT 44

PCT-US93-09945-1
; Sequence 1, Application PC/TUS9309945
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses Related
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09945
; FILING DATE:
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/991,997
FILING DATE: 17-DEC-1992
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13...888
PCT-US93-09945-1

Query Match 1.4%; Score 17; DB 5; Length 1089;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 949 CTGCCTACTCAAGGAGC 965
DB 168 CTGCCTACTCAAGGAGC 184

RESULT 45
US-09-566-921-66
Sequence 66, Application US/09566921
Patent No. 6682888
GENERAL INFORMATION:
APPLICANT: Loring, Jeanne F.
APPLICANT: Tingley, Debra M.
APPLICANT: Edwards, Carla M.
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
FILE REFERENCE: PA-0024 US
CURRENT APPLICATION NUMBER: US/09/566,921
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL Program
SEQ ID NO 66
LENGTH: 1747
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6682888 244561.6
US-09-566-921-66

Query Match 1.4%; Score 17; DB 4; Length 1747;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGAATCCCTAAGGAGCA 122
DB 303 AGAATCCCTAAGGAGCA 319

RESULT 46
US-08-765-889C-1
Sequence 1, Application US/08765889C
Patent No. 6136572
GENERAL INFORMATION:
APPLICANT: BENATTI, Luca
APPLICANT: BRETTON, Jerome
APPLICANT: SPECIALE, Carmela
APPLICANT: OKUNO, Etsuo
APPLICANT: SCHWARCZ, Robert
APPLICANT: MOSCA, Monica
TITLE OF INVENTION: RECOMBINANT KAT ENZYME AND
PROCESS FOR ITS PREPARATION
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,889C
FILING DATE: 23-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1748 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-765-889C-1

Query Match 1.4%; Score 17; DB 3; Length 1748;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 903 TCCAGGCCCTGTGGAC 919
DB 345 TCCAGGCCCTGTGGAC 361

RESULT 47
PCT-US95-07855-1
Sequence 1, Application PC/TUS9507855
GENERAL INFORMATION:
APPLICANT: BENATTI, Luca
APPLICANT: BRETTON, Jerome
APPLICANT: SPECIALE, Carmela
APPLICANT: OKUNO, Etsuo
APPLICANT: SCHWARCZ, Robert
APPLICANT: MOSCA, Monica
TITLE OF INVENTION: RECOMBINANT KAT ENZYME AND
PROCESS FOR ITS PREPARATION
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07855
FILING DATE: 23-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1748 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US95-07855-1

Query Match 1.4%; Score 17; DB 5; Length 1748;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 903 TCCAGGCCCTGGTGGAC 919
Db 345 TCCAGGCCCTGGTGGAC 361

RESULT 48
US-09-252-991A-6975
Sequence 6975, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6975
LENGTH: 1953
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6975

Query Match 1.4%; Score 17; DB 4; Length 1953;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 692 CTGAGCAGGAGAAAGC 708
Db 371 CTGAGCAGGAGAAAGC 387

RESULT 49
US-09-252-991A-6727/c
Sequence 6727, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6727
LENGTH: 2118
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6727

Query Match 1.4%; Score 17; DB 4; Length 2118;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 692 CTGAGCAGGAGAAAGC 708
Db 642 CTGAGCAGGAGAAAGC 626

RESULT 50
US-08-755-559-2/c
Sequence 2, Application US/08755559
Patent No. 5912142
GENERAL INFORMATION:
APPLICANT: KAUFMAN, RUSSEL E.
TITLE OF INVENTION: SLENTZ-KESLER, KIMBERLY
TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/755,559
FILING DATE: 22-NOV-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2180 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-755-559-2

Query Match 1.4%; Score 17; DB 2; Length 2180;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 272 GAAGGGCCCCAAAGCC 288
Db 1532 GAAGGGCCCCAAAGCC 1516

RESULT 51
US-09-210-474-2/c
Sequence 2, Application US/09210474
Patent No. 6072034
GENERAL INFORMATION:
APPLICANT: KAUFMAN, RUSSEL E.
TITLE OF INVENTION: SLENTZ-KESLER, KIMBERLY
TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.

STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,474
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/755,559
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2180 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-210-474-2

Query Match 1.4%; Score 17; DB 3; Length 2180;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 272 GAAGGGCCCCCAAGCC 288
Db 1532 GAAGGGCCCCCAAGCC 1516

RESULT 52
US-09-539-774-2/c
Sequence 2, Application US/09539774
Patent No. 6350615
GENERAL INFORMATION:
APPLICANT: KAUFMAN, RUSSEL E.
APPLICANT: SLENTZ-KESLER, KIMBERLY
TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
TITLE OF INVENTION: CELLS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/539,774
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/210,474
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.

REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2180 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-539-774-2

Query Match 1.4%; Score 17; DB 4; Length 2180;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 272 GAAGGGCCCCCAAGCC 288
Db 1532 GAAGGGCCCCCAAGCC 1516

RESULT 53
US-09-373-157-5
Sequence 5, Application US/09373157
Patent No. 6418963
GENERAL INFORMATION:
APPLICANT: GRIENINGER, Gerd
APPLICANT: APPELGADE, DIANE
APPLICANT: STOIKE-STEBEN, LARA
TITLE OF INVENTION: NOVEL CLEAVED FRAGMENTS OF FIBRINOGEN
FILE REFERENCE: Sequence ID No. 6418963. 1-7 for 454-24
Patent No. 6418963
CURRENT APPLICATION NUMBER: US/09/373,157
CURRENT FILING DATE: 1999-08-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 2648
TYPE: DNA
ORGANISM: Homo sapiens
US-09-373-157-5

Query Match 1.4%; Score 17; DB 4; Length 2648;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1144 CAGCCTGAATGACGAGG 1160
Db 2138 CAGCCTGAATGACGAGG 2154

RESULT 54
US-09-566-921-3/c
Sequence 3, Application US/09566921
Patent No. 6682888
GENERAL INFORMATION:
APPLICANT: LORING, Jeanne F.
APPLICANT: TINGLEY, Debra M.
APPLICANT: EDWARDS, Carla M.
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
FILE REFERENCE: PA-0024 US
CURRENT APPLICATION NUMBER: US/09/566,921
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL Program
SEQ ID NO 3
LENGTH: 2666
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. 6682888 232838.13
US-09-566-921-3

Query Match 1.4%; Score 17; DB 4; Length 2666;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 750 TCCTCATCCGGGAGAGC 766
Db 902 TCCTCATCCGGGAGAGC 886

RESULT 55

US-09-252-991A-6767/c
Sequence 6767, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6767
LENGTH: 2856
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6767

Query Match 1.4%; Score 17; DB 4; Length 2856;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 692 CTGAGCAGGAGAGAGC 708
Db 1394 CTGAGCAGGAGAGAGC 1378

RESULT 56

US-09-179-558-54
Sequence 54, Application US/09179558
Patent No. 6180612
GENERAL INFORMATION:
APPLICANT: Hockensmith, Joel W.
APPLICANT: Mathuswami, Rohini
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
TITLE OF INVENTION: TARGETING DNA METABOLIC PROCESSES USING
TITLE OF INVENTION: AMINOGLYCOSIDE DERIVATIVES
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/179,558
FILING DATE: 27-OCT-1998
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 09/060,470
FILING DATE: 15-APR-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 60/063,898
FILING DATE: 31-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 9426-005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 2874 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other
US-09-179-558-54

Query Match 1.4%; Score 17; DB 3; Length 2874;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 404 AGTCTGCCGAGAGAGC 420
Db 781 AGTCTGCCGAGAGAGC 797

RESULT 57

US-09-722-825-54
Sequence 54, Application US/09722825
Patent No. 6531306
GENERAL INFORMATION:
APPLICANT: Hockensmith, Joel W.
APPLICANT: Mathuswami, Rohini
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
TITLE OF INVENTION: TARGETING DNA METABOLIC PROCESSES USING
TITLE OF INVENTION: AMINOGLYCOSIDE DERIVATIVES
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/722,825
FILING DATE: 28-NO. 6531306-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/179,558
FILING DATE: <Unknown>
APPLICATION NUMBER: U.S. 60/063,898
FILING DATE: 31-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 9426-005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 2874 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-722-825-54

Query Match 1.4%; Score 17; DB 4; Length 2874;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 AGTCTGCCCGCAGAG 420
Db 781 AGTCTGCCCGCAGAG 797

RESULT 58
US-09-722-487-54
; Sequence 54, Application US/09722487
; Patent No. 6537791
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; Muthuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TARGETING DNA METABOLIC PROCESSES USING
; AMINOGLYCOSIDE DERIVATIVES

NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSER: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/722,487
FILING DATE: 28-No. 6537791-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/179,558
FILING DATE: <Unknown>
APPLICATION NUMBER: U.S. 60/063,898
FILING DATE: 31-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 9426-005-999

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:

LENGTH: 2874 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: Other
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-722-487-54

Query Match 1.4%; Score 17; DB 4; Length 2874;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 AGTCTGCCCGCAGAG 420
Db 781 AGTCTGCCCGCAGAG 797

RESULT 59
US-09-722-708-54
; Sequence 54, Application US/09722708
; Patent No. 6573060
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; Muthuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TARGETING DNA METABOLIC PROCESSES USING
; AMINOGLYCOSIDE DERIVATIVES

NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSER: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/722,708
FILING DATE: 28-No. 6573060-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/179,558
FILING DATE: <Unknown>
APPLICATION NUMBER: U.S. 60/063,898
FILING DATE: 31-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 9426-005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:

LENGTH: 2874 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-722-708-54

Query Match 1.4%; Score 17; DB 4; Length 2874;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 AGTCTGCCCGCAGAG 420
Db 781 AGTCTGCCCGCAGAG 797

RESULT 60
US-09-179-558-61
; Sequence 61, Application US/09179558
; Patent No. 6180612
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; Muthuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TARGETING DNA METABOLIC PROCESSES USING
; AMINOGLYCOSIDE DERIVATIVES

NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSER: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/722,487
FILING DATE: 28-No. 6537791-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/179,558
FILING DATE: <Unknown>
APPLICATION NUMBER: U.S. 60/063,898
FILING DATE: 31-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 9426-005-999

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:

LENGTH: 2874 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: Other
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-722-487-54

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/179,558
FILING DATE: 27-OCT-1998
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 09/060,470
FILING DATE: 15-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 60/063,898
FILING DATE: 31-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 9426-005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 3059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-179-558-61

Query Match 1.4%; Score 17; DB 3; Length 3059;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 404 AGTCTGCCCGACGAGAAG 420
|||||
Db 966 AGTCTGCCCGACGAGAAG 982

RESULT 61
US-09-722-825-61
; Sequence 61, Application US/09722825
; Patent No. 6531306
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; Muthuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TARGETING DNA METABOLIC PROCESSES USING
; AMINOGLYCOSIDE DERIVATIVES
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/722,825
; FILING DATE: 28-NO. 6531306-2000
; CLASSIFICATION: <Unknown>

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/179,558
FILING DATE: 31-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 9426-005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 3059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-722-825-61

Query Match 1.4%; Score 17; DB 4; Length 3059;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 404 AGTCTGCCCGACGAGAAG 420
|||||
Db 966 AGTCTGCCCGACGAGAAG 982

RESULT 62
US-09-722-487-61
; Sequence 61, Application US/09722487
; Patent No. 6537791
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; Muthuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TARGETING DNA METABOLIC PROCESSES USING
; AMINOGLYCOSIDE DERIVATIVES
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/722,487
; FILING DATE: 28-NO. 6537791-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/179,558
; FILING DATE: <Unknown>
; APPLICATION NUMBER: U.S. 60/063,898
; FILING DATE: 31-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 9426-005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)7909090
; TELEFAX: (212)8699741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 61:
; CLASSIFICATION: <Unknown>

```
SEQUENCE CHARACTERISTICS:
LENGTH: 3059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-722-487-61

Query Match      1.4%; Score 17; DB 4; Length 3059;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 AGCTGCCAGCAGAG 420
DB 966 AGCTGCCAGCAGAG 982

RESULT 63
US-09-722-708-61
; Sequence 61, Application US/09722708
; Patent No. 6573060
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; Mithuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TARGETING DNA METABOLIC PROCESSES USING
; AMINOGLYCOSIDE DERIVATIVES
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/722,708
; FILING DATE: 28-NOV-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/179,558
; FILING DATE: <Unknown>
; APPLICATION NUMBER: U.S. 60/063,898
; FILING DATE: 31-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 9426-005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)7909090
; TELEFAX: (212)8699741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3059 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-722-708-61

Query Match      1.4%; Score 17; DB 4; Length 3059;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 AGCTGCCAGCAGAG 420
```

```
DB 966 AGCTGCCAGCAGAG 982

RESULT 64
US-09-620-312D-349/c
; Sequence 349, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 349_
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (260)..(2164)
US-09-620-312D-349

Query Match      1.4%; Score 17; DB 4; Length 4139;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 499 GGCCACAGCCGTGGCC 515
DB 487 GGCCACAGCCGTGGCC 471

RESULT 65
US-08-896-449A-1
; Sequence 1, Application US/08896449A
; Patent No. 6040143
; GENERAL INFORMATION:
; APPLICANT: Vonta, Patrick J
; APPLICANT: Yuzbasiyan-Gurkan, Vilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
; TITLE OF INVENTION: FACTOR AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: 5445 Corporate Drive
; CITY: Troy
; STATE: Michigan
; COUNTRY: USA
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8802
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-132-652-1

Query Match      1.4%; Score 17; DB 3; Length 8802;
Best Local Similarity 100.0%; Pred. No. 1.12e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      136 CCATCCCTGGTGACAA 152
DB      6815 CCATCCCTGGTGACAA 6831

RESULT 67
US-09-534-638-1/C
; Sequence 1, Application US/09534638
; Patent No. 6320038
; GENERAL INFORMATION:
; APPLICANT: Panula, Pertti A. J.
; APPLICANT: Brandt, Annika
; APPLICANT: Westerlund, Johanna
; TITLE OF INVENTION: Promoter for Neuropeptide FF Promoter and use thereof
; TITLE OF INVENTION: for therapy and diagnosis
; FILE REFERENCE: 2530-104
; CURRENT APPLICATION NUMBER: US/09/534,638
; CURRENT FILING DATE: 2000-03-27
; EARLIER APPLICATION NUMBER: 09/365755
; EARLIER FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9840
; TYPE: DNA
; ORGANISM: Mouse
US-09-534-638-1

Query Match      1.4%; Score 17; DB 4; Length 9840;
Best Local Similarity 100.0%; Pred. No. 1.11e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      376 CTGAGTGCTCTGCTGAG 392
DB      1838 CTGAGTGCTCTGCTGAG 1822

RESULT 68
US-09-453-702B-62
; Sequence 62, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; APPLICANT: Burland, Valerie
; APPLICANT: Perna, Nicole T.
; APPLICANT: Plunkett, Guy
; APPLICANT: Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44MB storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:

```

```

; ZIP: 48098
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,449A
; FILING DATE: 18-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REFERENCE/DOCKET NUMBER: 2115-001226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 248-641-1600
; TELEFAX: 248-641-0270
; TELEX: 287637
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8802 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 203..8641
; OTHER INFORMATION: /function= "Blood Clotting Protein"
; OTHER INFORMATION: /product= "canine von Willebrand Factor"
; OTHER INFORMATION: /standard_name= "vWF"
; PUBLICATION INFORMATION:
; AUTHORS: Vonta, Patrick J.
; AUTHORS: Li, Jianping
; AUTHORS: Yuzbasivan-Gurkan, Vilma
; AUTHORS: Schall, William D.
; AUTHORS: Brewer, George J.
; TITLE: Von Willebrand's Disease in the Scottish
; TITLE: Terrier is Caused by a Single Base Deletion in
; TITLE: Exon Four of the von Willebrand Factor Gene
; JOURNAL: Journal of the American Veterinary Medicine Association
; DATE: 1996
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 8802
US-08-896-449A-1

Query Match      1.4%; Score 17; DB 3; Length 8802;
Best Local Similarity 100.0%; Pred. No. 1.12e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      136 CCATCCCTGGTGACAA 152
DB      6815 CCATCCCTGGTGACAA 6831

RESULT 66
US-09-132-652-1
; Sequence 1, Application US/09132652
; Patent No. 6074832
; GENERAL INFORMATION:
; APPLICANT: Vonta, Patrick J
; APPLICANT: Yuzbasivan-Gurkan, Vilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
; APPLICANT: Duffendeck, John
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 2115S-001226CPB
; CURRENT APPLICATION NUMBER: US/09/132,652
; CURRENT FILING DATE: 1998-08-11
; EARLIER APPLICATION NUMBER: 08/896,449
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 29

```

; APPLICATION NUMBER: US/09/453,702B
 ; FILING DATE: 03-Dec-1999
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/110,955
 ; FILING DATE: 04-Dec-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seay, Nicholas J.
 ; REGISTRATION NUMBER: 27386
 ; REFERENCE/DOCKET NUMBER: 960296.95017
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (608) 251-5000
 ; TELEFAX: (608) 251-9166
 ; INFORMATION FOR SEQ ID NO: 62:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 61663
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 62:
 ; US-09-453-702B-62

Query Match 1.4%; Score 17; DB 4; Length 61663;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 921 ATTACTCTGAGCTGCG 937
 DB 49896 ATTACTCTGAGCTGCG 49912

RESULT 69
 US-09-103-840A-2
 ; Sequence 2, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 4403765
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; FEATURE:
 ; OTHER INFORMATION: CDC 1551
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence
 ; OTHER INFORMATION: represent a, t, c or g
 ; US-09-103-840A-2

Query Match 1.4%; Score 17; DB 3; Length 4403765;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 AGCCATTGACCATCGTC 583
 DB 155 AGCCATTGACCATCGTC 171

RESULT 70
 US-09-103-840A-1
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.

; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 ; US-09-103-840A-1

Query Match 1.4%; Score 17; DB 3; Length 4411529;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 AGCCATTGACCATCGTC 583
 DB 155 AGCCATTGACCATCGTC 171

RESULT 71
 US-09-404-417A-8/c
 ; Sequence 8, Application US/09404417A
 ; Patent No. 6627729
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; APPLICANT: Deisher, Theresa A.
 ; APPLICANT: Jaspers, Stephen R.
 ; TITLE OF INVENTION: TML PEPTIDES
 ; FILE REFERENCE: 97-04C1
 ; CURRENT APPLICATION NUMBER: US/09/404,417A
 ; CURRENT FILING DATE: 1999-09-23
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 18
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: OLIGONUCLEOTIDE
 ; US-09-404-417A-8

Query Match 1.4%; Score 16; DB 4; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 GTCCAGCCAGGCGATG 311
 DB 16 GTCCAGCCAGGCGATG 1

RESULT 72
 US-09-833-381-378/c
 ; Sequence 378, Application US/09833381
 ; Patent No. 6672186
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; CURRENT FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 378
 ; LENGTH: 211


```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-378

Query Match      1.4%; Score 16; DB 4; Length 211;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      628 CAGAGAGTATAACATC 643
Db      120 CAGAGAGTATAACATC 105

RESULT 73
US-09-602-877A-93
; Sequence 93, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446CS
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-93

Query Match      1.4%; Score 16; DB 4; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      24 GCCTGTGTCTCTGTGA 39
Db      92 GCCTGTGTCTCTGTGA 107

RESULT 74
US-09-071-710-9/c
; Sequence 9, Application US/09071710
; Patent No. 6130043
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/525,397
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/071,710
; FILING DATE:

Query Match      1.4%; Score 16; DB 3; Length 265;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      24 GCCTGTGTCTCTGTGA 39
Db      141 GCCTGTGTCTCTGTGA 126

RESULT 75
US-09-525-397-9/c
; Sequence 9, Application US/09525397
; Patent No. 6252047
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/525,397
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/071,710
; FILING DATE:

```

ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 265 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-525-397-9

Query Match 1.4%; Score 16; DB 3; Length 265;
Best Local Similarity 100.0%; Pred.No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GCCTGTGCTCTGTGA 39
Db 141 GCCTGTGCTCTGTGA 126

RESULT 76

US-09-621-976-16442
Sequence 16442, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054P2

CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm
SEQ ID NO 16442
LENGTH: 278

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 114
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16442

Query Match 1.4%; Score 16; DB 4; Length 278;
Best Local Similarity 100.0%; Pred.No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 739 TGGAGGGGCGCTTCCTC 754
Db 230 TGGAGGGGCGCTTCCTC 245

RESULT 77

US-09-313-294A-3534/c
Sequence 3534, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:

APPLICANT: Ialugudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PI-0017 US

CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600

SOFTWARE: PERL Program
SEQ ID NO 3534

LENGTH: 283
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6476212 700611809HI
US-09-313-294A-3534

Query Match 1.4%; Score 16; DB 4; Length 283;
Best Local Similarity 100.0%; Pred.No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 163 GCACACAGATGCTGAG 178
Db 176 GCACACAGATGCTGAG 161

RESULT 78

US-09-071-710-10/c
Sequence 10, Application US/09071710
Patent No. 6130043
GENERAL INFORMATION:

APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.

APPLICANT: HODGES, STEVEN C.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.

APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road

CITY: Abbott Park
STATE: IL
COUNTRY: USA

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,710
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/850,713
FILING DATE: 02-MAY-1997

ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6083.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

FEATURE:
NAME/KEY: base_polymorphism
LOCATION: 147

OTHER INFORMATION: /note= " N' represents an A or G or
OTHER INFORMATION: T or C polymorphism at this position"
US-09-071-710-10

Query Match 1.4%; Score 16; DB 3; Length 288;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GCCTGTGTCTCTGTGA 39
Db 62 GCCTGTGTCTCTGTGA 47

RESULT 79

US-09-525-397-10/c
Sequence 10, Application US/09525397
Patent No. 6252047
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/525.397
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071,710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: base polymorphism
LOCATION: 147
OTHER INFORMATION: /note= " N' represents an A or G or
OTHER INFORMATION: T or C polymorphism at this position"
US-09-525-397-10

Query Match 1.4%; Score 16; DB 3; Length 288;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GCCTGTGTCTCTGTGA 39
Db 62 GCCTGTGTCTCTGTGA 47

RESULT 80

US-08-235-838-9/c
Sequence 9, Application US/08235838
Patent No. 5571894
GENERAL INFORMATION:
APPLICANT: Wels, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5571894man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
Growth Factor Receptor
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,838
FILING DATE: TBA
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: . PMW15-VL51-1
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..310
OTHER INFORMATION: /note= "1-18 partial seq. of
VK1BACK primer region; 64-96 CDR1L; 142-162 CDR2L;
OTHER INFORMATION: 259-282 CDR3L; 292-310 partial seq. of VK1FOR
OTHER INFORMATION: primer region

US-08-235-838-9

Query Match 1.4%; Score 16; DB 1; Length 310;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 586 TGAGGATGGAGACTGG 601
Db 24 TGAGGATGGAGACTGG 9

RESULT 81

US-08-465-473B-9/c
Sequence 9, Application US/08465473B

Patent No. 5939531

GENERAL INFORMATION:

APPLICANT: Wells, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5939531man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
Growth Factor Receptor
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: NOVARTIS Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901-6940

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,473B
FILING DATE: 5 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pfeiffer, Heena J.
REGISTRATION NUMBER: 22,640
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908)522 6940
TELEFAX: (908)522 6955
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
INDIVIDUAL ISOLATE: B. coli
IMMEDIATE SOURCE:
CLONE: pW15-VL51-1
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..310

OTHER INFORMATION: /note="1-18 partial seq. of

VK1BACK primer region; 64-96 CDR1L; 142-162 CDR2L;

OTHER INFORMATION: 259-282 CDR3L; 292-310 partial seq. of VK1FOR
OTHER INFORMATION: Primer region
US-08-465-473B-9

Query Match 1.4%; Score 16; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 586 TGAGGATGGAGACTGG 601
Db 24 TGAGGATGGAGACTGG 9

RESULT 82

US-08-888-366-21/c
Sequence 21, Application US/08888366

Patent No. 5972656

GENERAL INFORMATION:

APPLICANT: Lopez, Osvaldo
APPLICANT: Wylie, Dwane E.
APPLICANT: Wagner, Fred W.
TITLE OF INVENTION: Mercury Binding Polypeptides and Nucleotides Coding Therefore
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 90 South 7th Street, 3100 No. 5972656west Ctr.
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/888,366
FILING DATE: 03-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/187,407
FILING DATE: 27-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,542
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/493,299
FILING DATE: 14-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/324,392
FILING DATE: 14-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Carter, Charles G.
REGISTRATION NUMBER: 35,093
REFERENCE/DOCKET NUMBER: 8648.39USC1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
STRAIN: Light chain variable region for monoclonal
antibody 23F8
FEATURE:
NAME/KEY: CDS
LOCATION: 1..321

US-08-888-366-21

```
Query Match      1.4%; Score 16; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      586 TGAGATGGAGACTGG 601
Db      30 TGAGATGGAGACTGG 15

RESULT 83
US-09-641-638-433/c
; Sequence 433, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 433
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 250
; OTHER INFORMATION: 10-40-252 : polymorphic base C or T
; NAME/KEY: misc_binding
; LOCATION: 230..249
; OTHER INFORMATION: 10-40-252.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 251..270
; OTHER INFORMATION: 10-40-252.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 1..18
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 335..352
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 238..262
; OTHER INFORMATION: 10-40-252 potential probe
; US-09-641-638-434

Query Match      1.4%; Score 16; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      896 CCCTCACTCCAGGCC 911
Db      207 CCCTCACTCCAGGCC 192

RESULT 85
US-09-621-976-18404/c
; Sequence 18404, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PE2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18404
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-621-976-18404

Query Match      1.4%; Score 16; DB 4; Length 391;

Qy      896 CCCTCACTCCAGGCC 911
Db      207 CCCTCACTCCAGGCC 192

RESULT 84
US-09-641-638-434/c
; Sequence 434, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
```

Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1132 CAGCTTCTACATCAGC 1147
DB 294 CAGCTTCTACATCAGC 279

Query Match 1.4%; Score 16; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 86
US-09-489-039A-3397
; Sequence 3397, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Berton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3397
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3397

Query Match 1.4%; Score 16; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GGCCGAGCTGTCGCTG 556
DB 105 GGCCGAGCTGTCGCTG 120

RESULT 87
US-09-564-329A-10/c
; Sequence 10, Application US/09564329A
; Patent No. 6541212
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; APPLICANT: Safir, Douglas C.
; TITLE OF INVENTION: PROSTATE STEM CELL ANTIGEN AND USES THEREOF
; FILE REFERENCE: 30435.54US14
; CURRENT APPLICATION NUMBER: US/09/564,329A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 09/359,326
; PRIOR FILING DATE: 1998-07-20
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 60/113,230
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/120,536
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 60/124,658
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 09/203,939
; PRIOR FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: 09/251,835
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 09/308,503
; PRIOR FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10
; LENGTH: 408
; TYPE: DNA
; ORGANISM: SCID Mice
US-09-564-329A-10

Query Match 1.4%; Score 16; DB 4; Length 408;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 CACTCCAGGCCCTGGT 915
DB 179 CACTCCAGGCCCTGGT 164

RESULT 88
US-08-822-028-62
; Sequence 62, Application US/08822028
; Patent No. 5993813
; GENERAL INFORMATION:
; APPLICANT: MEZES, PETER S
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: ANDERSON, WH KERR
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHOLON, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
; MODIFIED ANTIBODIES FOR CANCER TREATMENT
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DUANE C ULMER
; STREET: P.O. BOX 1967
; CITY: MIDLAND
; STATE: MICHIGAN
; COUNTRY: USA
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/040,687
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-37,075C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-08-822-028-62

Query Match 1.4%; Score 16; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 CTGACTGCAGACAGAT 172
DB 283 CTGACTGCAGACAGAT 298

RESULT 89
US-08-822-028-62
; Sequence 62, Application US/08822028
; Patent No. 5993813
; GENERAL INFORMATION:
; APPLICANT: MEZES, PETER S
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: ANDERSON, WH KERR
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHOLON, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
; MODIFIED ANTIBODIES FOR CANCER TREATMENT
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DUANE C ULMER
; STREET: P.O. BOX 1967
; CITY: MIDLAND
; STATE: MICHIGAN
; COUNTRY: USA
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/040,687
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-37,075C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-08-822-028-62

Query Match 1.4%; Score 16; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 CTGACTGCAGACAGAT 172
DB 283 CTGACTGCAGACAGAT 298

APPLICATION NUMBER: US 08/155,301
FILING DATE: 18-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,739
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/165,699
FILING DATE: 10-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,741
FILING DATE: 09-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/728,463
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/16433
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,417
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/21803
FILING DATE: 01-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 014643-009040US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 360:
SEQUENCE CHARACTERISTICS:
LENGTH: 439 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-042-353-360

Query Match 1.4%; Score 16; DB 3; Length 439;
Best Local Similarity 100.0%; Pred.No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 586 TGAGGATGGAGACTGG 601
|||
Db 96 TGAGGATGGAGACTGG 81

RESULT 92

US-08-758-417A-208/c
Sequence 208, Application US/08758417A
Patent No. 6300129
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
Kay, Robert M.

TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for
Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 417
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,417A
FILING DATE: 02-DEC-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/728,463
FILING DATE: 10-OCT-1996
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
APPLICATION NUMBER: US 08/209,741
FILING DATE: 09-MAR-1994
APPLICATION NUMBER: US 08/165,699
FILING DATE: 10-DEC-1993
APPLICATION NUMBER: US 08/161,739
FILING DATE: 03-DEC-1993
APPLICATION NUMBER: US 08/155,301
FILING DATE: 18-NOV-1993
APPLICATION NUMBER: US 08/096,762
FILING DATE: 22-JUL-1993
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Serafini, Andrew T.
REGISTRATION NUMBER: 41,303
REFERENCE/DOCKET NUMBER: 014643-009030US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 439 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 208:
US-08-758-417A-208

Query Match 1.4%; Score 16; DB 4; Length 439;
Best Local Similarity 100.0%; Pred.No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 586 TGAGGATGGAGACTGG 601
|||
Db 96 TGAGGATGGAGACTGG 81

RESULT 93

US-09-621-976-10604
Sequence 10604, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 10604
LENGTH: 480
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:


```
; NAME/KEY: misc_feature
; LOCATION: 474
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-10604

Query Match      1.4%; Score 16; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1012 CCTACCTGTGACTGTG 1027
Db 106 CCTACCTGTGACTGTG 121
|||||

RESULT 94
US-09-621-976-2590/c
; Sequence 2590, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2590
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 214...456
US-09-621-976-2590

Query Match      1.4%; Score 16; DB 4; Length 497;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 966 CCTGTGTCTCTCAGAG 981
Db 281 CCGTGTCTCTCAGAG 266
|||||

RESULT 95
US-09-252-991A-1532
; Sequence 1532, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1532
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1532

Query Match      1.4%; Score 16; DB 4; Length 546;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 906 AGGCCCTGTGTGACCA 921
|||||

US-09-939-853a-74.olg.rni

Db 379 AGGCCCTGTGTGACCA 394
|||||

RESULT 96
US-09-328-352-491/c
; Sequence 491, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 491
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-491

Query Match      1.4%; Score 16; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 TCTGCTGAGGACAAT 399
Db 229 TCTGCTGAGGACAAT 214
|||||

RESULT 97
US-09-252-991A-11119
; Sequence 11119, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11119
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11119

Query Match      1.4%; Score 16; DB 4; Length 594;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 GCCGTGGCCCTGGGCA 521
Db 349 GCCGTGGCCCTGGGCA 364
|||||

RESULT 98
US-09-252-991A-12620
; Sequence 12620, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
```

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12620
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12620

Query Match 1.4%; Score 16; DB 4; Length 615;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1057 GGACAGCTCCCTCTG 1072
DB 234 GGACAGCTCCCTCTG 249

RESULT 99
US-09-252-991A-9984/c
; Sequence 9984 Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9984
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9984

Query Match 1.4%; Score 16; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 TGTCCAAGGCCAGGGA 469
DB 270 TGTCCAAGGCCAGGGA 255

RESULT 100
US-09-252-991A-11241
; Sequence 11241 Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11241
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11241

Query Match 1.4%; Score 16; DB 4; Length 654;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 506 GCCGTGGCCCTGGGCA 521
DB 206 GCCGTGGCCCTGGGCA 221

Search completed: March 25, 2004, 07:40:20
Job time : 121 secs

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 07:38:49 ; Search time 456 Seconds
(without alignments)
9658.598 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctcaaggacc.....tctcttgatgatgcctag 1183

Scoring table: OLIGO NUC

Gapop 60.0, Gapext 60.0

Searched: 2458946 seqs, 1861504846 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10A_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1183	100.0	1183	12 US-09-939-853A-74
2	1183	100.0	1183	12 US-09-939-853A-74
3	735	62.1	786	14 US-10-043-649-1
4	724	61.2	864	10 US-09-814-353-21302
5	657	55.5	763	9 US-09-867-550-953
6	348	29.4	444	9 US-09-867-550-951
7	341	28.8	875	9 US-09-867-550-1915
8	134	11.3	432	9 US-09-864-761-2829
9	134	11.3	448	9 US-09-864-761-15513
10	96	8.1	96	9 US-09-864-761-13612
11	87	7.4	320	10 US-09-814-353-17314
12	77	6.5	152	10 US-09-814-353-4631
13	77	6.5	152	10 US-09-814-353-10930
14	26	2.2	26	12 US-09-939-853A-141
15	22	1.9	22	12 US-09-939-853A-142
16	1183	100.0	1183	12 US-09-939-853A-74
17	1183	100.0	1183	12 US-09-939-853A-74
18	735	62.1	786	14 US-10-043-649-1
19	724	61.2	864	10 US-09-814-353-21302
20	657	55.5	763	9 US-09-867-550-953
21	348	29.4	444	9 US-09-867-550-951
22	341	28.8	875	9 US-09-867-550-1915
23	134	11.3	432	9 US-09-864-761-2829
24	134	11.3	448	9 US-09-864-761-15513
25	96	8.1	96	9 US-09-864-761-13612
26	87	7.4	320	10 US-09-814-353-17314
27	77	6.5	152	10 US-09-814-353-4631
28	77	6.5	152	10 US-09-814-353-10930
29	26	2.2	26	12 US-09-939-853A-141
30	22	1.9	22	12 US-09-939-853A-142

Sequence 2188, Ap	US-10-062-674-2188	701	1.8	21	C 16
Sequence 140, App	US-09-939-853A-140	20	1.7	20	C 17
Sequence 195852,	US-10-027-633-195852	12	1.7	12	C 18
Sequence 107077,	US-10-027-633-107077	611	1.7	611	C 19
Sequence 142058,	US-10-027-633-142058	672	1.7	672	C 20
Sequence 143059,	US-10-027-633-143059	672	1.7	672	C 21
Sequence 142060,	US-10-027-633-142060	672	1.7	672	C 22
Sequence 26286, A	US-10-027-633-26286	711	1.7	711	C 23
Sequence 640, App	US-10-027-633-640	934	1.7	934	C 24
Sequence 1776, Ap	US-10-063-674-1776	2017	1.7	2017	C 25
Sequence 27, Appl	US-10-366-288-27	2032	1.7	2032	C 26
Sequence 954, App	US-09-960-708-954	2129	1.7	2129	C 27
Sequence 1452, Ap	US-10-303-720-1452	2129	1.7	2129	C 28
Sequence 3, Appl	US-10-413-277-3	786431	1.7	786431	C 29
Sequence 4510, Ap	US-09-908-975-4510	65	1.6	65	C 30
Sequence 30106, A	US-09-864-761-30106	114	1.6	114	C 31
Sequence 237, App	US-10-073-602B-237	510	1.6	510	C 32
Sequence 13565, A	US-09-864-761-13565	559	1.6	559	C 33
Sequence 20, Appl	US-09-789-561-20	1033	1.6	1033	C 34
Sequence 118578,	US-10-027-633-118578	1123	1.6	1123	C 35
Sequence 5547, A	US-10-424-599-55347	1133	1.6	1133	C 36
Sequence 11, Appl	US-10-354-358-11	1467	1.6	1467	C 37
Sequence 25, Appl	US-10-280-576-25	1490	1.6	1490	C 38
Sequence 25, Appl	US-09-976-783-25	1580	1.6	1580	C 39
Sequence 1, Appl	US-09-861-846-1	1833	1.6	1833	C 40
Sequence 577, App	US-10-094-749-577	2120	1.6	2120	C 41
Sequence 1101, Ap	US-10-305-720-1101	2771	1.6	2771	C 42
Sequence 46381, A	US-10-369-493-46381	4207	1.6	4207	C 43
Sequence 2048, Ap	US-10-062-674-2048	4720	1.6	4720	C 44
Sequence 3261, Ap	US-09-738-626-3261	204	1.5	204	C 45
Sequence 8609, Ap	US-09-918-995-8609	402	1.5	402	C 46
Sequence 1815, Ap	US-09-983-965-1815	403	1.5	403	C 47
Sequence 195591,	US-10-027-633-195591	443	1.5	443	C 48
Sequence 26739, A	US-09-918-995-26739	487	1.5	487	C 49
Sequence 38019, A	US-09-918-995-38019	491	1.5	491	C 50
Sequence 270409,	US-10-027-633-270409	497	1.5	497	C 51
Sequence 284851,	US-10-027-633-284851	498	1.5	498	C 52
Sequence 284852,	US-10-027-633-284852	498	1.5	498	C 53
Sequence 143161,	US-10-027-633-143161	531	1.5	531	C 54
Sequence 143162,	US-10-027-633-143162	531	1.5	531	C 55
Sequence 4315, Ap	US-10-029-386-4315	535	1.5	535	C 56
Sequence 86, Appl	US-09-764-853-86	544	1.5	544	C 57
Sequence 282391,	US-10-027-633-282391	561	1.5	561	C 58
Sequence 72725, A	US-10-424-599-72725	578	1.5	578	C 59
Sequence 6891, Ap	US-10-029-386-6891	584	1.5	584	C 60
Sequence 264852,	US-10-027-633-264852	608	1.5	608	C 61
Sequence 162462,	US-10-027-633-162462	717	1.5	717	C 62
Sequence 144544,	US-10-027-633-144544	718	1.5	718	C 63
Sequence 144545,	US-10-027-633-144545	718	1.5	718	C 64
Sequence 325220,	US-10-027-633-325220	823	1.5	823	C 65
Sequence 325447,	US-10-027-633-325447	833	1.5	833	C 66
Sequence 172936,	US-10-027-633-172936	858	1.5	858	C 67
Sequence 66566, A	US-10-424-599-66566	944	1.5	944	C 68
Sequence 120486,	US-10-027-633-120486	999	1.5	999	C 69
Sequence 216, App	US-10-120-988-216	1194	1.5	1194	C 70
Sequence 28115, A	US-10-425-114-28115	1232	1.5	1232	C 71
Sequence 1261, Ap	US-10-260-238-1261	1325	1.5	1325	C 72
Sequence 4, Appl	US-09-997-701-4	1438	1.5	1438	C 73
Sequence 8, Appl	US-09-931-836-8	1579	1.5	1579	C 74
Sequence 441, App	US-10-206-915-441	1579	1.5	1579	C 75
Sequence 441, App	US-10-199-670-441	1579	1.5	1579	C 76
Sequence 441, App	US-10-201-858-441	1579	1.5	1579	C 77
Sequence 441, App	US-10-205-890-441	1579	1.5	1579	C 78
Sequence 441, App	US-10-208-024-441	1579	1.5	1579	C 79
Sequence 441, App	US-10-201-853-441	1579	1.5	1579	C 80
Sequence 8, Appl	US-10-036-342-8	1579	1.5	1579	C 81
Sequence 441, App	US-10-052-586-441	1579	1.5	1579	C 82
Sequence 8, Appl	US-10-036-041-8	1579	1.5	1579	C 83
Sequence 8, Appl	US-10-035-855-8	1579	1.5	1579	C 84
Sequence 441, App	US-10-174-590-441	1579	1.5	1579	C 85
Sequence 441, App	US-10-176-758-441	1579	1.5	1579	C 86
Sequence 441, App	US-10-175-737-441	1579	1.5	1579	C 87
Sequence 441, App	US-10-173-706-441	1579	1.5	1579	C 88


```

Db      661 CAAAGTCTCCATGGGTGGCTGTATGAGGCGCTGAGCAGGGAGAAAGCAGAGAACTGCT 720
Qy      721 GTTGTATCCTGGGAACCCCTGGAGGGCCCTTCCTCATCCGGGAGAGCCAGCAGAGAGAG 780
Db      721 GTTGTATCCTGGGAACCCCTGGAGGGCCCTTCCTCATCCGGGAGAGCCAGCAGAGAGAG 780
Qy      781 CTCTTACTCTCTGTCTGAGTCCGCTCAGCGCCCTGATCTCTGAGCTGGGATGACACTA 840
Db      781 CTCTTACTCTCTGTCTGAGTCCGCTCAGCGCCCTGATCTCTGAGCTGGGATGACACTA 840
Qy      841 CAGGATCCACTGCTTGTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACTTCCCTC 900
Db      841 CAGGATCCACTGCTTGTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACTTCCCTC 900
Qy      901 ACTCCAGGCGCTGGTGGACATTAATCTCTGAGCTGGCGGATGACATCTGCTGCTACTCA 960
Db      901 ACTCCAGGCGCTGGTGGACATTAATCTCTGAGCTGGCGGATGACATCTGCTGCTACTCA 960
Qy      961 GGAGCCCTGTCTCTCAGAGGGCTGGCGCGCTCCCTGGCAGGATATACCCCTACTGT 1020
Db      961 GGAGCCCTGTCTCTCAGAGGGCTGGCGCGCTCCCTGGCAGGATATACCCCTACTGT 1020
Qy      1021 GACTGTGAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTCTGA 1080
Db      1021 GACTGTGAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTCTGA 1080
Qy      1081 AGCTGCCAGAGGAGGAGTCTCTCTGAGTGGGCTGGCGCGCTCCCTGGCAGGATATAC 1140
Db      1081 AGCTGCCAGAGGAGGAGTCTCTCTGAGTGGGCTGGCGCGCTCCCTGGCAGGATATAC 1140
Qy      1141 CATCAGCTGGAATGAGGAGGCTGTCTCTTTGGATGATGCTAG 1183
Db      1141 CATCAGCTGGAATGAGGAGGCTGTCTCTTTGGATGATGCTAG 1183

```

RESULT 2

```

US-09-939-853A-76/c
; Sequence 76, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-939-853A-76

```

```

Query Match      100.08; Score 1183; DB 12; Length 1183;
Best Local Similarity 100.08; Pred. No. 0;
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AGCTAGAGCTCCAGAGGCCACCCCTGCTGTCTCTGTGACAGAGCTCAAGGGGCCCTGGG 60
Db      1183 AGCTAGAGCTCCAGAGGCCACCCCTGCTGTCTCTGTGACAGAGCTCAAGGGGCCCTGGG 1124

Qy      61 CCTTCCCTCCCTGGCTGGCTGGTGGAGGGTCCCGAGTCCAGATCCCTTAAGGAG 120
Db      1123 CCTTCCCTCCCTGGCTGGCTGGTGGAGGGTCCCGAGTCCAGATCCCTTAAGGAG 1064

```

```

Qy      121 CATGGGCGAGTGATCCATCCCTGGTGTACAAATGCTGTGACTGCAGACAGATGCTGAGCT 180
Db      1063 CATGGGCGAGTGATCCATCCCTGGTGTACAAATGCTGTGACTGCAGACAGATGCTGAGCT 1004

Qy      181 ACCCAAAACCAACCTAGCTCTCCCTGGAAGATCTCCAGGCTGAGAGAGTCTGGGTG 240
Db      1003 ACCCAAAACCAACCTAGCTCTCCCTGGAAGATCTCCAGGCTGAGAGAGTCTGGGTG 944

Qy      241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAAAGGGCCCCCAAGCCCTAACTGTCCA 300
Db      943 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAAAGGGCCCCCAAGCCCTAACTGTCCA 884

Qy      301 GCGAGAGATGGCTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAATTTCC 360
Db      883 GCGAGAGATGGCTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAATTTCC 824

Qy      361 CTCGATCATGTCTGTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTCCAGCAGAGAAG 420
Db      823 CTCGATCATGTCTGTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTCCAGCAGAGAAG 764

Qy      421 AAAATCTCTCCAAAGCCCAAGCTTGAATTCCTCTGTCCAAGGCCAGGAGCTGTGACCAAT 480
Db      763 AAAATCTCTCCAAAGCCCAAGCTTGAATTCCTCTGTCCAAGGCCAGGAGCTGTGACCAAT 704

Qy      481 GGAAGCAGAGAGAGCAAGGCCACAGCGCTGGCCCTGGGAGTTCCTCCGCGAGTGGCCCC 540
Db      703 GGAAGCAGAGAGAGCAAGGCCACAGCGCTGGCCCTGGGAGTTCCTCCGCGAGTGGCCCC 644

Qy      541 GGCAGAGCTGTCTGAGACTTCGGGAGGCAATGACCATCGTCTCTGAGGATGGAGACTG 600
Db      643 GGCAGAGCTGTCTGAGACTTCGGGAGGCAATGACCATCGTCTCTGAGGATGGAGACTG 584

Qy      601 GTGGAAGGCTGTCTGAGTCTCAGCAGAGAGATTAACATCCCGAGCGTCCAGCTGGG 660
Db      583 GTGGAAGGCTGTCTGAGTCTCAGCAGAGAGATTAACATCCCGAGCGTCCAGCTGGG 524

Qy      661 CAAAGTCTCCATGGGTGGCTGTATGAGGCGCTGAGCAGGGAGAAAGCAGAGAACTGCT 720
Db      523 CAAAGTCTCCATGGGTGGCTGTATGAGGCGCTGAGCAGGGAGAAAGCAGAGAACTGCT 464

Qy      721 GTTGTATCCTGGGAACCCCTGGAGGGCCCTTCCTCATCCGGGAGAGCCAGCAGAGAGG 780
Db      463 GTTGTATCCTGGGAACCCCTGGAGGGCCCTTCCTCATCCGGGAGAGCCAGCAGAGAGG 404

Qy      781 CTCTTACTCTCTGTCTGAGTCCGCTCAGCGCCCTGCATCTCTGGGACCGGATCAGACACTA 840
Db      403 CTCTTACTCTCTGTCTGAGTCCGCTCAGCGCCCTGCATCTCTGGGACCGGATCAGACACTA 344

Qy      841 CAGGATCCACTGCTTGTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTC 900
Db      343 CAGGATCCACTGCTTGTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTC 284

Qy      901 ACTCCAGGCGCTGGTGGACATTAATCTCTGAGCTGGCGGATGACATCTGCTGCTACTCA 960
Db      283 ACTCCAGGCGCTGGTGGACATTAATCTCTGAGCTGGCGGATGACATCTGCTGCTACTCA 224

Qy      961 GGAGCCCTGTCTCTCAGAGGGCTGGCGCGCTCCCTGGCAGGATATACCCCTACTGT 1020
Db      223 GGAGCCCTGTCTCTCAGAGGGCTGGCGCGCTCCCTGGCAGGATATACCCCTACTGT 164

Qy      1021 GACTGTGAGAGGACACCACTCAATGGAAAGAGCTGGACAGCTCCCTCTCTGTTTCTGA 1080
Db      163 GACTGTGAGAGGACACCACTCAATGGAAAGAGCTGGACAGCTCCCTCTCTGTTTCTGA 104

Qy      1081 AGCTGCCAGAGGAGGAGTCTCTTCTCAGTGGGGTCTCCGGGAGTCCCTCAGCTTCTA 1140
Db      103 AGCTGCCAGAGGAGGAGTCTCTTCTCAGTGGGGTCTCCGGGAGTCCCTCAGCTTCTA 44

Qy      1141 CATCAGCTGGAATGAGGAGGCTGTCTCTTTGGATGATGCTAG 1183
Db      43 CATCAGCTGGAATGAGGAGGCTGTCTCTTTGGATGATGCTAG 1

```

878 TCACCGCGCTCACCTTCCCTCCTCCTCAGGCGCTGGTGGACCATCTCTCAGCTGGG 937
481 TCACCGCGCTCACCTTCCCTCCTCCTCAGGCGCTGGTGGACCATCTCTCAGCTGGG 540
938 GATGACATCTGCTGCTCCTACTCAAGAGCGCTGTCTCTGACAGGCGCTGCGCTCCCT 997
541 GATGACATCTGCTGCTCCTACTCAAGAGCGCTGTCTCTGACAGGCGCTGCGCTCCCT 600
998 GCAAGGATATACCCCTACCTGTGCTGACAGGAGACACCACTCACTGGAAGAGCTG 1057
601 GCAAGGATATACCCCTACCTGTGCTGACAGGAGACACCACTCACTGGAAGAGCTG 660
1058 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTCTCAGTGAAGGT 1117
661 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTCTCAGTGAAGGT 720
1118 CTCGGGAGTCCCTCAGCTTCTATCAGCTGCAATGACAGGCGTCTCTCTGAGTAT 1177
721 CTCGGGAGTCCCTCAGCTTCTATCAGCTGCAATGACAGGCGTCTCTCTGAGTAT 780
1178 GCCTAG 1183
781 GCCTAG 786

RESULT 4
US-09-814-353-21302
; Sequence 21302, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MEI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21302
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 32, 862, 863, 864
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21302

Query Match 61.2%; Score 724; DB 10; Length 864;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTAGAGCTCCAGGAGCCCGCTGCTGTCTGTGACAGAGCTCAAGGGCCCTGGGC 61
DB 54 GCTAGAGCTCCAGGAGCCCGCTGCTGTCTGTGACAGAGCTCAAGGGCCCTGGGC 113
QY 62 CTCCTCCCTCCCTGGCTCGGCTGTGCTGTGGAGGGTTCGCCAGTCCAGAAATCCCTAAGGAGC 121

US-10-043-649-1
; Sequence 1, Application US/10043649
; Publication No. US2003005924A1
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha J.
; APPLICANT: Mendenhall, Marcy K.
; APPLICANT: Pardo, Jorge
; APPLICANT: Spencer, Collin
; APPLICANT: Fu, C. Alan
; APPLICANT: Luo, Ying
; APPLICANT: Payan, Donald G.
; APPLICANT: Mancebo, Helena S.Y.
; APPLICANT: Wu, Jun
; APPLICANT: Zhou, Xiulan
; APPLICANT: Shen, Mary
; APPLICANT: Liao, X. Charlene
; APPLICANT: Sheng, Ning
; TITLE OF INVENTION: Cloning of a No. US2003005924A1el Inhibitor of Antigen-receptor
; FILE REFERENCE: A-70219-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/10/043,649
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/260,953
; PRIOR FILING DATE: 2001-01-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(786)
; OTHER INFORMATION:
US-10-043-649-1

Query Match 62.1%; Score 735; DB 14; Length 786;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGGAAGTCTGCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC 457
DB 1 ATGGGAAGTCTGCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC 60
QY 458 CAAGGCCCGGACCTGTGACCTGAAGCAGAGAGAGAGGACAGGCGCTGGCCCTG 517
DB 61 CAAGGCCCGGACCTGTGACCTGAAGCAGAGAGAGAGGACAGGCGCTGGCCCTG 120
QY 518 GGCAGTTTCCCGGAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 577
DB 121 GGCAGTTTCCCGGAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
QY 578 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGTGAAGTCTCAGGAGAGATAT 637
DB 181 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGTGAAGTCTCAGGAGAGATAT 240
QY 638 AACATCCCGAGCGTCCAGTGGGCAAGTCCCATGGTGGTGGTGTATGAGGCGCTGAGC 697
DB 241 AACATCCCGAGCGTCCAGTGGGCAAGTCCCATGGTGGTGGTGTATGAGGCGCTGAGC 300
QY 698 AGGAGAGAGCAGAGAACTGCTGTTTACCTGGGAAACCCCTGGAGGGGCGCTTCTCATC 757
DB 301 AGGAGAGAGCAGAGAACTGCTGTTTACCTGGGAAACCCCTGGAGGGGCGCTTCTCATC 360
QY 758 CGGAGAGCAGAGCAGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCGGCTCAGCGGCTGCA 817
DB 361 CGGAGAGCAGAGCAGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCGGCTCAGCGGCTGCA 420
QY 818 TCTCGGAGCGGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACATC 877
DB 421 TCTCGGAGCGGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACATC 480


```

; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 951
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-550-951

Query Match      29.4%; Score 348; DB 9; Length 444;
Best Local Similarity 100.0%; Pred. No. 7.9e-167;
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 CCTAGGACCAAGGACACTGGCAGAGCTTCCAGAGGGGCCCAAGCCCTAACCTGTCCAG 301
DB 1 CCTAGGACCAAGGACACTGGCAGAGCTTCCAGAGGGGCCCAAGCCCTAACCTGTCCAG 60

QY 302 CCAGAGCATGGCTCTCAGCAGAGCTGTCTTCCAGAGCTTGTGATGACAAACCAATTTCCC 361
DB 61 CCAGAGCATGGCTCTCAGCAGAGCTGTCTTCCAGAGCTTGTGATGACAAACCAATTTCCC 120

QY 362 TCGATGATGTCTTCTGAGTGTCTGCTGAGGAAACATGGAAAGTCTGCCAGCAGAGA 421
DB 121 TCGATGATGTCTTCTGAGTGTCTGCTGAGGAAACATGGAAAGTCTGCCAGCAGAGA 180

QY 422 AAATCTTCCAAAGCCCAAGCTTGAATTCCTCTGTCCAGGCCAGGACCTGTGACCATG 481
DB 181 AAATCTTCCAAAGCCCAAGCTTGAATTCCTCTGTCCAGGCCAGGACCTGTGACCATG 240

QY 482 GAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
DB 241 GAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

QY 542 GCCGAGCTGTGCTGAGAGCTCGGGAGGACCATGACCATGCTCTCTGAG 589
DB 301 GCCGAGCTGTGCTGAGAGCTCGGGAGGACCATGACCATGCTCTCTGAG 348

RESULT 7
US-09-867-550-1915
; Sequence 1915, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1915
; LENGTH: 875
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: Wherein n is one of a or t or c or g
US-09-867-550-1915

Query Match      28.8%; Score 341; DB 9; Length 875;
Best Local Similarity 100.0%; Pred. No. 2.7e-163;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 843 GGATCCACTGCTTGACAAATGGCTGGCTGTATCTTCCAGGGGCTCAGCTTCCCTCAG 902

```

```

DB 2 GGATCCACTGCTTGACAAATGGCTGGCTGTATCTTCCAGGGGCTCAGCTTCCCTCAG 61
QY 903 TCCAGGCCCTGCTGGACCAATTAATCTGAGCTGGCGGATGACATCTGCTTACTCAAGG 962
DB 62 TCCAGGCCCTGCTGGACCAATTAATCTGAGCTGGCGGATGACATCTGCTTACTCAAGG 121
QY 963 AGCCCTGTGTCTGTCAGAGGGCTGGCCCGCTCCCTGGGAGGATATACCCCTACCTGTGA 1022
DB 122 AGCCCTGTGTCTGTCAGAGGGCTGGCCCGCTCCCTGGGAGGATATACCCCTACCTGTGA 181
QY 1023 CTGTGTCAGAGGACACCACTCAACTGGAAGAGCTGGAGAGCTCCCTCTGTTTCTGTAAG 1082
DB 182 CTGTGTCAGAGGACACCACTCAACTGGAAGAGCTGGAGAGCTCCCTCTGTTTCTGTAAG 241
QY 1083 CTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACA 1142
DB 242 CTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACA 301
QY 1143 TCAGCCCTGAATGACGAGGCTGTCTCTTCTGATGATGCTAG 1183
DB 302 TCAGCCCTGAATGACGAGGCTGTCTCTTCTGATGATGCTAG 342

RESULT 8
US-09-864-761-2829
; Sequence 2829, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,697
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

```



```
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19612
; LENGTH: 96
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031662.24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: NT HIT: AF000716.1, EVALUE 1.70e-01
; OTHER INFORMATION: EST_HUMAN HIT: A1125308.1, EVALUE 2.10e-01
US-09-864-761-19612

Query Match      8.1%; Score 96; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.5e-38;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 967 CTGTGCTCTGAGAGGGTGGCCGCTCCCTGGCAAGGATATACCCCTACTGTGACTGT 1026
Db 1 CTGTGCTCTGAGAGGGTGGCCGCTCCCTGGCAAGGATATACCCCTACTGTGACTGT 60

Qy 1027 GCAGAGGACACCACTCACTGGAAGAGCTGGACAG 1062
Db 61 GCAGAGGACACCACTCACTGGAAGAGCTGGACAG 96

RESULT 11
US-09-814-353-17314
; Sequence 17314, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4631
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 17, 102, 112
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-4631

Query Match      6.5%; Score 77; DB 10; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.1e-28;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 549 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGTGTGACGG 608
Db 103 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGTGTGACGG 162

Qy 609 TGCTGTCTGAAGTCTCAGGCAGAGAGT 635
Db 163 TGCTGTCTGAAGTCTCAGGCAGAGAGT 189

RESULT 12
US-09-814-353-4631
; Sequence 4631, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4631
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 17, 102, 112
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-4631

Query Match      6.5%; Score 77; DB 10; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.1e-28;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 549 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGTGTGACGG 608
Db 103 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAGATGAGACTGTGTGACGG 162
```

Db 25 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTGTGAGGATGGAGACTGGTGGACGG 84

QY 609 TGCTGTCTGAAGTCTCA 625
Db 85 TGCTGTCTGAAGTCTCA 101

RESULT 13

US-09-814-353-10930
; Sequence 10930, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-0068
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10930
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 17, 102, 112
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-10930

Query Match 6.5%; Score 77; DB 10; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.1e-28;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTGTGAGGATGGAGACTGGTGGACGG 608
Db 25 TGTGCTGAGACTCGGGAGCCATTGACCATGCTCTGTGAGGATGGAGACTGGTGGACGG 84

QY 609 TGCTGTCTGAAGTCTCA 625
Db 85 TGCTGTCTGAAGTCTCA 101

RESULT 14

US-09-939-853A-141/c
; Sequence 141, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961

; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 141
; LENGTH: 26
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide primer
US-09-939-853A-141

Query Match 2.2%; Score 26; DB 12; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 AAGGACACTGGCAGACTTCCAGAAG 276
Db 26 AAGGACACTGGCAGACTTCCAGAAG 1

RESULT 15

US-09-939-853A-142
; Sequence 142, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 142
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide primer
US-09-939-853A-142

Query Match 1.9%; Score 22; DB 12; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 TGAGAGAGTTCTGGGTGTCCTA 245
Db 1 TGAGAGAGTTCTGGGTGTCCTA 22

RESULT 16

US-10-062-674-2188/c
; Sequence 2188, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24

NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 2188
; LENGTH: 701
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040005559A1 893157.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (701)
; OTHER INFORMATION: a, t, c, g, or other
; OTHER INFORMATION: 2188
US-10-062-674-2188
Query Match 1.8%; Score 21; DB 15; Length 701;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 510 TGGCCCTGGGCGAGTTCCCGG 530
Db 285 TGGCCCTGGGCGAGTTCCCGG 265
RESULT 17
US-09-939-853A-140/C
; Sequence 140, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR FILING DATE: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 140
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide primer
US-09-939-853A-140
Query Match 1.7%; Score 20; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 282 CAAGCCCTACCTGTCAG 301
Db 20 CAAGCCCTACCTGTCAG 1
RESULT 18
US-10-027-632-195852/c
; Sequence 195852, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195852
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-195852
Query Match 1.7%; Score 20; DB 15; Length 611;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 256 CACTGGCAGACTTCCAGAG 275
Db 503 CACTGGCAGACTTCCAGAG 484
RESULT 19
US-10-027-632-107077
; Sequence 107077, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107077
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-107077
Query Match 1.7%; Score 20; DB 15; Length 672;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 747 CCTTCCTCATCCGGAGAC 766
Db 71 CCTTCCTCATCCGGAGAC 90

```
RESULT 20
US-10-027-632-142058
; Sequence 142058, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142058
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142058

Query Match          1.7%; Score 20; DB 15; Length 672;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGAC 766
DB 71 CCTTCTCATCCGGGAGAC 90

RESULT 21
US-10-027-632-142059
; Sequence 142059, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142059
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142059

Query Match          1.7%; Score 20; DB 15; Length 672;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGAC 766
DB 71 CCTTCTCATCCGGGAGAC 90
```

```
US-10-027-632-142059

Query Match          1.7%; Score 20; DB 15; Length 672;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGAC 766
DB 71 CCTTCTCATCCGGGAGAC 90

RESULT 22
US-10-027-632-142060
; Sequence 142060, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142060
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142060

Query Match          1.7%; Score 20; DB 15; Length 672;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGAC 766
DB 71 CCTTCTCATCCGGGAGAC 90

RESULT 23
US-10-027-632-26286
; Sequence 26286, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
```

Db 421 CCGCGCGAGCTCTCGCTGA 441

RESULT 25

US-10-062-674-1776
; Sequence 1776, Application US/1006266
; Publication No. US2004005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser
; TITLE OF INVENTION: MARKERS OF NEURONS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/000000

; PRIOR APPLICATION NUMBER: US 09/625,102
 ;
 ; PRIOR FILING DATE: 2000-07-24
 ;
 ; NUMBER OF SEQ ID NOS: 2217
 ;

```

; SEQ ID NO 1776
; LENGTH: 2017
; TYPE: DNA

```

```

; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040005559A1 245648.12
US-10-062-674-1776

```

```

Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      747 CCTTCCTCATCCGGGAGAGC 756
        |||||
Db      549 CCTTCCTCATCCGGGAGAGC 568

RESULT 26
US-10-366-288-27
; Sequence 27, Application US/10366288

```

Publication No. US20030216288A1

```

? GENERAL INFORMATION:
? APPLICANT: Powell, Douglas
? APPLICANT: Welch, Nadine S.
? TITLE OF INVENTION: AIDS AND HIV-RELATED DISORDERS USING
? TITLE OF INVENTION: 1414, 1481, 1553,
? TITLE OF INVENTION: 34081, 1720, 1683, 1552, 1562, 1675, 12825, 9952, 5816,
? TITLE OF INVENTION: 10002, 1611, 1371, 44324, 126, 270, 312, 167, 326, 18926
? TITLE OF INVENTION: 6747, 1793, 1784 OR 2045 MOLECULES
? FILE REFERENCE: MP102-025P1RNMNM
? CURRENT APPLICATION NUMBER: US/10/366,288
?

```

```

, CURRENT FILING DATE: 2003-02-13
, PRIOR APPLICATION NUMBER: 60/357,391
, PRIOR FILING DATE: 2002-02-15
, PRIOR APPLICATION NUMBER: 60/380,249
, PRIOR FILING DATE: 2002-05-13
, PRIOR APPLICATION NUMBER: 60/391,306
, PRIOR FILING DATE: 2002-06-25
, PRIOR APPLICATION NUMBER: 60/406,297
, PRIOR FILING DATE: 2002-08-27
, PRIOR APPLICATION NUMBER: 60/412,007
, PRIOR FILING DATE: 2002-09-19
, PRIOR APPLICATION NUMBER: 60/417,508
, PRIOR FILING DATE: 2002-10-10
, PRIOR APPLICATION NUMBER: 60/432,318
, PRIOR FILING DATE: 2002-12-10
, NUMBER OF SEQ IDS NOS: 52
, SOFTWARE: SeqSEQ for Windows Version 4.0
,

```

```

LENGTH: 2032
TYPE: DNA
ORGANISM: Homo Sapien

```

Query Match 1.7%; Score 20; DB 15; Length 2032;

```
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGAC 766
Db 500 CCTTCCTCATCCGGGAGAC 519

RESULT 27
US-09-960-706-954
; Sequence 954, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 954
; LENGTH: 2129
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 U23852
US-09-960-706-954

Query Match 1.7%; Score 20; DB 10; Length 2129;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGAC 766
Db 508 CCTTCCTCATCCGGGAGAC 527

RESULT 28
US-10-305-720-1452
; Sequence 1452, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1452
; LENGTH: 2129
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g775207
US-10-305-720-1452

Query Match 1.7%; Score 20; DB 15; Length 2129;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGAC 766
Db 508 CCTTCCTCATCCGGGAGAC 527
```

```
RESULT 29
US-10-412-277-3
; Sequence 3, Application US/10412277
; Publication No. US20030175791A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001067DIV
; CURRENT APPLICATION NUMBER: US/10/412,277
; CURRENT FILING DATE: 2003-04-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(786431)
; OTHER INFORMATION: n = A,T,C or G
US-10-412-277-3

Query Match 1.7%; Score 20; DB 14; Length 786431;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 TGAGCAGGGGAGAAAGCAGAG 712
Db 412751 TGAGCAGGGGAGAAAGCAGAG 412770

RESULT 30
US-09-908-975-4510
; Sequence 4510, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE V
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4510
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-908-975-4510

Query Match 1.6%; Score 19; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 395 ACAATGGGAAGTCTGCCCA 413
Db 2 ACAATGGGAAGTCTGCCCA 20

RESULT 31
US-09-864-761-30106/C
; Sequence 30106, Application US/09864761
```


Publication No. US20030109670A1

GENERAL INFORMATION:

APPLICANT: University of Utah Research Foundation

APPLICANT: Cognetix, Inc.

APPLICANT: Olivera, Baldomero M.

APPLICANT: McIntosh, J, Michael

APPLICANT: Watkins, Maren

APPLICANT: Garrett, James E.

APPLICANT: Cruz, Lourdes J.

APPLICANT: Grilley, Michelle

APPLICANT: Schoenfeld, Robert M.

APPLICANT: Walker, Craig

APPLICANT: Shetty, Reshma

APPLICANT: Jones, Robert M.

TITLE OF INVENTION: Cone Snail Peptides

FILE REFERENCE: 2314-249

CURRENT APPLICATION NUMBER: US/10/072,602B

CURRENT FILING DATE: 2002-02-11

PRIOR APPLICATION NUMBER: US 60/267,408

PRIOR FILING DATE: 2001-02-09

NUMBER OF SEQ ID NOS: 638

SOFTWARE: PatentIn version 3.0

SEQ ID NO 237

LENGTH: 510

TYPE: DNA

ORGANISM: Conus textile

FEATURE:

NAME/KEY: CDS

LOCATION: (223)..(471)

US-10-072-602B-237

Query Match 1.6%; Score 19; DB 14; Length 510;

Best Local Similarity 100.0%; Pred. No. 34;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GGAAACAATGGGAAGTCTGC 410

|||||

DB 390 GGAAACAATGGGAAGTCTGC 408

|||||

RESULT 33

US-09-864-761-13565/c

Sequence 13565, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aecomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/006666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006665

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 13565
LENGTH: 599
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC020596.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
US-09-864-761-13565

Query Match 1.6%; Score 19; DB 9; Length 599;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 751 CCTCATCCGGGAGAGCCAG 769
Db 75 CCTCATCCGGGAGAGCCAG 57

RESULT 34
US-09-789-561-20/c
Sequence 20, Application US/09789561
Patent No. US20020064818A1
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: 52 Human secreted proteins
FILE REFERENCE: P2043P1
CURRENT FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: PCT/US00/24008
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 60/152,317
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/152,315
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 1033
TYPE: DNA
ORGANISM: Homo sapiens
US-09-789-561-20

Query Match 1.6%; Score 19; DB 9; Length 1033;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 GCAGGAGAGAGAGAGCA 714
Db 931 GCAGGAGAGAGAGAGCA 913

RESULT 35
US-10-027-632-118578/c
Sequence 118578, Application US/10027632

Publication No. US20030204075A9
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 118578

LENGTH: 1125

TYPE: DNA

ORGANISM: Human

US-10-027-632-118578

Query Match 1.6%; Score 19; DB 15; Length 1125;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 964 GCCCTGTGTCTGCAGAGG 982
Db 926 GCCCTGTGTCTGCAGAGG 908

RESULT 36
US-10-424-599-55347
Sequence 55347, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 55347
LENGTH: 1133
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_20990C.1
US-10-424-599-55347

Query Match 1.6%; Score 19; DB 12; Length 1133;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 TCCCCAGCAGAGAGAAATC 426
Db 834 TCCCCAGCAGAGAGAAATC 852

RESULT 37
US-10-354-358-11

```

; Sequence 11, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 9990, 2100, 9288, 64588, 10480, 20893, 33230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6176, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MP102-020P1RNMNM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1467)
; US-10-354-358-11

```

```

Query Match 1.6%; Score 19; DB 14; Length 1467;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 742 AGGGGCTTCCTCATCCGG 760
    |||||
Db 423 AGGGGCTTCCTCATCCGG 441
    |||||

```

```

RESULT 38
US-10-280-576-25
; Sequence 25, Application US/10280576
; Publication No. US20040044405A1
; GENERAL INFORMATION:
; APPLICANT: Wolff, Matthew R.
; TITLE OF INVENTION: VASCULAR STENT OR GRAFT COATED OR IMPREGNATED WITH PROTEIN
; FILE REFERENCE: 09820.189
; CURRENT APPLICATION NUMBER: US/10/280,576
; CURRENT FILING DATE: 2002-10-22

```

```

; PRIOR APPLICATION NUMBER: 60/343,732
; PRIOR FILING DATE: 2001-10-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-280-576-25

Query Match 1.6%; Score 19; DB 12; Length 1490;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 AGGGGCTTCCTCATCCGG 760
    |||||
Db 434 AGGGGCTTCCTCATCCGG 452
    |||||

RESULT 39
US-09-976-782-25
; Sequence 25, Application US/09976782
; Publication No. US20030190715A1
; GENERAL INFORMATION:
; APPLICANT: Grosse et al
; TITLE OF INVENTION: No. US20030190715A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-157
; CURRENT APPLICATION NUMBER: US/09/976,782
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,113
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,662
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,732
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,625
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,703
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/241,190
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,637
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,669
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/262,455
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/240,648
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 1580
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-976-782-25

```

```

Query Match 1.6%; Score 19; DB 10; Length 1580;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 742 AGGGGCTTCCTCATCCGG 760
    |||||
Db 455 AGGGGCTTCCTCATCCGG 473
    |||||

```

```

RESULT 40
US-09-861-846-1
; Sequence 1, Application US/09861846
; Patent No. US20020110852A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

```

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001065
; CURRENT APPLICATION NUMBER: US/09/861,846
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 09/752,821
; PRIOR FILING DATE: 2001-01-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1833
; TYPE: DNA
; ORGANISM: Human
US-09-861-846-1

Query Match 1.6%; Score 19; DB 9; Length 1833;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTCATCC 758
Db 157 GGAGGGGCTTCTCATCC 175

RESULT 41

US-10-094-749-577
; Sequence 577, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 577
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-577

Query Match 1.6%; Score 19; DB 15; Length 2120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTCATCC 758
Db 280 GGAGGGGCTTCTCATCC 298

RESULT 42

US-10-062-674-2048/c
; Sequence 2048, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674

US-10-305-720-1101
; Sequence 1101, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1101
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 gi2556002
US-10-305-720-1101

Query Match 1.6%; Score 19; DB 15; Length 2771;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 585 CTGAGGATGGAGACTGGTG 603
Db 1305 CTGAGGATGGAGACTGGTG 1323

RESULT 43

US-10-369-493-46381/c
; Sequence 46381, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46381
; LENGTH: 4207
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-46381

Query Match 1.8%; Score 19; DB 15; Length 4207;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 CCAGCAGAGAAATCTCT 429
Db 3095 CCAGCAGAGAAATCTCT 3077

RESULT 44

US-10-062-674-2048/c
; Sequence 2048, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674

; CURRENT FILING DATE: 2002-01-30
 ; PRIOR APPLICATION NUMBER: US 09/625,102
 ; PRIOR FILING DATE: 2000-07-24
 ; NUMBER OF SEQ ID NOS: 2217
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 2048
 ; LENGTH: 4720
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20040005559A1 422072.14
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1) ... (4720)
 ; OTHER INFORMATION: a, t, c, g, or other
 US-10-062-674-2048

Query Match 1.6%; Score 19; DB 15; Length 4720;
 Best Local Similarity 100.0%; Pred. No. 25; Mismatches 0; Indels 0; Gaps 0;

Qy 585 CTGAGGATGGAGCTGGTG 603
 Db 3421 CTGAGGATGGAGCTGGTG 3403

RESULT 45

US-09-738-626-3261
 ; Sequence 3261, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: Patent in ver. 3.0
 ; SEQ ID NO 3261
 ; LENGTH: 204
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3261

Query Match 1.5%; Score 18; DB 9; Length 204;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 GAATGACGAGGTGCTC 1167
 Db 111 GAATGACGAGGTGCTC 128

RESULT 46

US-09-918-995-8609
 ; Sequence 8609, Application US/09918995
 ; Publication No. US20030073623A1

; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FROM VARIOUS CDNA LIBRARIES
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8609
 ; LENGTH: 402
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-918-995-8609

Query Match 1.5%; Score 18; DB 10; Length 402;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 416 AGAAGAAATCTCTGCCA 433
 Db 3 AGAAGAAATCTCTGCCA 20

RESULT 47

US-09-983-965-1815
 ; Sequence 1815, Application US/09983965
 ; Patent No. US20020137160A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.
 ; APPLICANT: Tao, Nengbing
 ; APPLICANT: Byatt, John C.
 ; APPLICANT: Mathialagan, Nagappan
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 ; FILE REFERENCE: 37-21(10297)C
 ; CURRENT APPLICATION NUMBER: US/09/983,965
 ; CURRENT FILING DATE: 2001-10-26
 ; PRIOR APPLICATION NUMBER: US 09/465,231
 ; PRIOR FILING DATE: 1999-12-15
 ; PRIOR APPLICATION NUMBER: US 60/113,678
 ; PRIOR FILING DATE: 1998-12-17
 ; NUMBER OF SEQ ID NOS: 5912
 ; SEQ ID NO 1815
 ; LENGTH: 403
 ; TYPE: DNA
 ; ORGANISM: Bos taurus
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 02-LIB3057-001-Q1-K1-A9
 US-09-983-965-1815

Query Match 1.5%; Score 18; DB 9; Length 403;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 902 CTCACGCCCTGGTGGAC 919
 Db 376 CTCACGCCCTGGTGGAC 393

RESULT 48

US-10-027-632-195991
 ; Sequence 195991, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195991
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-195991

Query Match 1.5%; Score 18; DB 15; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GCGCGAGCTGCTGCTGAG 558
DB 286 GCGCGAGCTGCTGCTGAG 303

RESULT 49

US-09-918-995-26739
; Sequence 26739, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26739
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-26739

Query Match 1.5%; Score 18; DB 10; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 ACGTGGCAAGTCTCCC 671
DB 448 ACGTGGCAAGTCTCCC 465

RESULT 50

US-09-918-995-38019
; Sequence 38019, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38019
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(491)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-38019

Query Match 1.5%; Score 18; DB 10; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 AGAAGAAATCTCTGCCA 433
DB 416 AGAAGAAATCTCTGCCA 433

RESULT 51

US-10-027-632-270409
; Sequence 270409, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 270409
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-270409

Query Match 1.5%; Score 18; DB 15; Length 497;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 CCAGAGGGCCCCCAAG 286
DB 277 CCAGAGGGCCCCCAAG 294

RESULT 52

US-10-027-632-284851/c
; Sequence 284851, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129

[illegible]

<pre>US-10-027-632-143161/c ; Sequence 143161, Application US/10027632 ; Publication No. US20030204075A9 ; GENERAL INFORMATION: ; APPLICANT: Wang, David G. ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide ; FILE REFERENCE: 108827.129 ; CURRENT APPLICATION NUMBER: US 10/027,632 ; PRIOR FILING DATE: 2002-04-30 ; PRIOR APPLICATION NUMBER: US 60/218,006 ; PRIOR FILING DATE: 2000-07-12 ; PRIOR APPLICATION NUMBER: US 60/198,676 ; PRIOR FILING DATE: 2000-04-20 ; PRIOR APPLICATION NUMBER: US 60/193,483 ; PRIOR FILING DATE: 2000-03-29 ; PRIOR APPLICATION NUMBER: US 60/185,218 ; PRIOR FILING DATE: 2000-02-24 ; PRIOR APPLICATION NUMBER: US 60/167,363 ; PRIOR FILING DATE: 1999-11-23 ; PRIOR APPLICATION NUMBER: US 60/156,358 ; PRIOR FILING DATE: 1999-09-28 ; PRIOR APPLICATION NUMBER: US 60/146,002 ; PRIOR FILING DATE: 1999-08-09 ; NUMBER OF SEQ ID NOS: 325720 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 143161 ; LENGTH: 531 ; TYPE: DNA ; ORGANISM: Human US-10-027-632-143161</pre>	<pre>Query Match 1.5%; Score 18; DB 15; Length 531; Best Local Similarity 100.0%; Pred.No. 1.le+02; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</pre>	
<pre>QY 587 GAGGATGGAGCTGTGG 604 Db 405 GAGGATGGAGCTGTGG 388</pre>	<pre>RESULT 53 US-10-027-632-284852/c ; Sequence 284852, Application US/10027632 ; Publication No. US20030204075A9 ; GENERAL INFORMATION: ; APPLICANT: Wang, David G. ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide ; FILE REFERENCE: 108827.129 ; CURRENT APPLICATION NUMBER: US 10/027,632 ; PRIOR FILING DATE: 2002-04-30 ; PRIOR APPLICATION NUMBER: US 60/218,006 ; PRIOR FILING DATE: 2000-07-12 ; PRIOR APPLICATION NUMBER: US 60/198,676 ; PRIOR FILING DATE: 2000-04-20 ; PRIOR APPLICATION NUMBER: US 60/193,483 ; PRIOR FILING DATE: 2000-03-29 ; PRIOR APPLICATION NUMBER: US 60/185,218 ; PRIOR FILING DATE: 2000-02-24 ; PRIOR APPLICATION NUMBER: US 60/167,363 ; PRIOR FILING DATE: 1999-11-23 ; PRIOR APPLICATION NUMBER: US 60/156,358 ; PRIOR FILING DATE: 1999-09-28 ; PRIOR APPLICATION NUMBER: US 60/146,002 ; PRIOR FILING DATE: 1999-08-09 ; NUMBER OF SEQ ID NOS: 325720 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 284851 ; LENGTH: 498 ; TYPE: DNA ; ORGANISM: Human US-10-027-632-284851</pre>	<pre>Query Match 1.5%; Score 18; DB 15; Length 498; Best Local Similarity 100.0%; Pred.No. 1.le+02; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</pre>
<pre>QY 587 GAGGATGGAGCTGTGG 604 Db 405 GAGGATGGAGCTGTGG 388</pre>	<pre>RESULT 53 US-10-027-632-284852/c ; Sequence 284852, Application US/10027632 ; Publication No. US20030204075A9 ; GENERAL INFORMATION: ; APPLICANT: Wang, David G. ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide ; FILE REFERENCE: 108827.129 ; CURRENT APPLICATION NUMBER: US 10/027,632 ; PRIOR FILING DATE: 2002-04-30 ; PRIOR APPLICATION NUMBER: US 60/218,006 ; PRIOR FILING DATE: 2000-07-12 ; PRIOR APPLICATION NUMBER: US 60/198,676 ; PRIOR FILING DATE: 2000-04-20 ; PRIOR APPLICATION NUMBER: US 60/193,483 ; PRIOR FILING DATE: 2000-03-29 ; PRIOR APPLICATION NUMBER: US 60/185,218 ; PRIOR FILING DATE: 2000-02-24 ; PRIOR APPLICATION NUMBER: US 60/167,363 ; PRIOR FILING DATE: 1999-11-23 ; PRIOR APPLICATION NUMBER: US 60/156,358 ; PRIOR FILING DATE: 1999-09-28 ; PRIOR APPLICATION NUMBER: US 60/146,002 ; PRIOR FILING DATE: 1999-08-09 ; NUMBER OF SEQ ID NOS: 325720 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 284852 ; LENGTH: 498 ; TYPE: DNA ; ORGANISM: Human US-10-027-632-284852</pre>	<pre>Query Match 1.5%; Score 18; DB 15; Length 498; Best Local Similarity 100.0%; Pred.No. 1.le+02; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</pre>
<pre>QY 587 GAGGATGGAGCTGTGG 604 Db 405 GAGGATGGAGCTGTGG 388</pre>	<pre>RESULT 53 US-10-027-632-284852/c ; Sequence 284852, Application US/10027632 ; Publication No. US20030204075A9 ; GENERAL INFORMATION: ; APPLICANT: Wang, David G. ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide ; FILE REFERENCE: 108827.129 ; CURRENT APPLICATION NUMBER: US 10/027,632 ; PRIOR FILING DATE: 2002-04-30 ; PRIOR APPLICATION NUMBER: US 60/218,006 ; PRIOR FILING DATE: 2000-07-12 ; PRIOR APPLICATION NUMBER: US 60/198,676 ; PRIOR FILING DATE: 2000-04-20 ; PRIOR APPLICATION NUMBER: US 60/193,483 ; PRIOR FILING DATE: 2000-03-29 ; PRIOR APPLICATION NUMBER: US 60/185,218 ; PRIOR FILING DATE: 2000-02-24 ; PRIOR APPLICATION NUMBER: US 60/167,363 ; PRIOR FILING DATE: 1999-11-23 ; PRIOR APPLICATION NUMBER: US 60/156,358 ; PRIOR FILING DATE: 1999-09-28 ; PRIOR APPLICATION NUMBER: US 60/146,002 ; PRIOR FILING DATE: 1999-08-09 ; NUMBER OF SEQ ID NOS: 325720 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 284852 ; LENGTH: 498 ; TYPE: DNA ; ORGANISM: Human US-10-027-632-284852</pre>	<pre>Query Match 1.5%; Score 18; DB 15; Length 498; Best Local Similarity 100.0%; Pred.No. 1.le+02; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</pre>

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-143162

Query Match
Best Local Similarity 1.5%; Score 18; DB 15; Length 531;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 897 CCTCACTCCAGGCCCTGG 914
|||||
DB 367 CCTCACTCCAGGCCCTGG 350

RESULT 56
US-10-029-386-4315
; Sequence 4315, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4315
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR22 175,0
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.46
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: B1518449.1, EVALUE 4.00e-89
; OTHER INFORMATION: NT HIT: g114779626, EVALUE 2.00e-89
; OTHER INFORMATION: SWISSPROT HIT: O00268, EVALUE 2.00e-01
US-10-029-386-4315

Query Match
Best Local Similarity 1.5%; Score 18; DB 14; Length 535;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 897 CCTCACTCCAGGCCCTGG 914
|||||
DB 135 CCTCACTCCAGGCCCTGG 152

RESULT 57
US-09-764-853-86
; Sequence 86, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 86
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (176)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (177)
```

```
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (190)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (500)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (522)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-853-86

Query Match
Best Local Similarity 1.5%; Score 18; DB 9; Length 544;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
|||||
DB 462 TACATCAGCCTGAATGAC 479

RESULT 58
US-10-027-632-282391/c
; Sequence 282391, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282391
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282391

Query Match
Best Local Similarity 1.5%; Score 18; DB 15; Length 561;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 TGAAGTCTCAGGCAGAGA 633
|||||
DB 114 TGAAGTCTCAGGCAGAGA 97

RESULT 59
US-10-424-599-72725/c
; Sequence 72725, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David X
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```


; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 72725
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36686C.1
US-10-424-599-72725

Query Match 1.5%; Score 18; DB 12; Length 578;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 388 CTGAGGAACAATGGGAAG 405
Db 105 CTGAGGAACAATGGGAAG 88

RESULT 60
US-10-029-386-6891/c
; Sequence 6891, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6891
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AB023048.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: AF168055.1, EVALUATE 1.40e-01
; OTHER INFORMATION: SWISSPROT HIT: P39605, EVALUATE 7.40e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA584104.1, EVALUATE 2.00e-04
US-10-029-386-6891

Query Match 1.5%; Score 18; DB 14; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 893 TTCCCTCCTCACTCCAGGCC 910
Db 130 TTCCCTCCTCACTCCAGGCC 113

RESULT 61
US-10-027-632-264852
; Sequence 264852, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264852
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-264852

Query Match 1.5%; Score 18; DB 15; Length 608;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 896 CCCTCACTCCAGCCCTG 913
Db 538 CCCTCACTCCAGCCCTG 555

RESULT 62
US-10-027-632-162462/c
; Sequence 162462, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162462
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162462

Query Match 1.5%; Score 18; DB 15; Length 717;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 885 GCCTCACCTTCCCTCAC 902
Db 155 GCCTCACCTTCCCTCAC 138

RESULT 63
US-10-027-632-144544/c
; Sequence 144544, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144544
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-144544

Query Match 1.5%; Score 18; DB 15; Length 718;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 729 CTGGGAACCTCGAGGGG 746
Db 289 CTGGGAACCTCGAGGGG 272

RESULT 64
US-10-027-632-144545/c
; Sequence 144545, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144545
; LENGTH: 718
; TYPE: DNA

US-10-027-632-144545
Query Match 1.5%; Score 18; DB 15; Length 718;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 729 CTGGGAACCTCGAGGGG 746
Db 289 CTGGGAACCTCGAGGGG 272

RESULT 65
US-10-027-632-325220/c
; Sequence 325220, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 325220
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-325220

Query Match 1.5%; Score 18; DB 15; Length 823;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 897 CCTCACTCCAGGCCCTGG 914
Db 418 CCTCACTCCAGGCCCTGG 401

RESULT 66
US-10-027-632-325447/c
; Sequence 325447, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24

;; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;; FILE REFERENCE: 38-21(53223)B
;; CURRENT APPLICATION NUMBER: US/10/424,599
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 265684
;; SEQ ID NO 66566
;; LENGTH: 944
;; TYPE: DNA
;; ORGANISM: Glycine max
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT3847_31122C.1
US-10-424-599-66566

Query Match 1.5%; Score 18; DB 12; Length 944;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 ACAAACTGCTGACTGCAG 166
DB 376 ACAAACTGCTGACTGCAG 393

RESULT 69
US-10-027-632-120486/c
; Sequence 120486, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120486
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(999)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-120486

Query Match 1.5%; Score 18; DB 15; Length 999;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 464 CAGGACCTGTGACCATG 481
DB 639 CAGGACCTGTGACCATG 622

RESULT 70
US-10-120-988-216
; Sequence 216, Application US/10120988
; Publication No. US20030219745A1

;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 325447
;; LENGTH: 823
;; TYPE: DNA
;; ORGANISM: Human
;; OTHER INFORMATION: Clone ID: PAT_MRT3847_31122C.1
US-10-027-632-325447

Query Match 1.5%; Score 18; DB 15; Length 823;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 897 CCTCACTCCAGCCCTGG 914
DB 418 CCTCACTCCAGCCCTGG 401

RESULT 67
US-10-027-632-172936/c
; Sequence 172936, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172936
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(999)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-172936

Query Match 1.5%; Score 18; DB 15; Length 858;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GGCTGTGCTTGGAGGGT 95
DB 344 GGCTGTGCTTGGAGGGT 327

RESULT 68
US-10-424-599-66566
; Sequence 66566, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Ren, Feiyun
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 802CON
 ; CURRENT APPLICATION NUMBER: US/10/120,988
 ; CURRENT FILING DATE: 2002-04-11
 ; PRIOR APPLICATION NUMBER: 09/774,528
 ; PRIOR FILING DATE: 2001-01-30
 ; NUMBER OF SEQ ID NOS: 441
 ; SOFTWARE: pt_FL_genes Version 2.0
 ; SEQ ID NO 216
 ; LENGTH: 1194
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(648)
 ; US-10-120-988-216

Query Match 1.5%; Score 18; DB 15; Length 1194;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GCGCGAGCTGTCGCTGAG 558
 DB 111 GCGCGAGCTGTCGCTGAG 128

RESULT 71
 US-10-425-114-28115/G
 ; Sequence 28115, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 28115
 ; LENGTH: 1232
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB4746-011-C3_FLI
 ; US-10-425-114-28115

Query Match 1.5%; Score 18; DB 12; Length 1232;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 TCCCTCGATGATGTCCTT 375
 DB 388 TCCCTCGATGATGTCCTT 371

RESULT 72
 US-10-260-238-1261
 ; Sequence 1261, Application US/10260238
 ; Publication No. US20040016028A1

; APPLICANT: Budworth, Paul R.
 ; APPLICANT: Moughamer, Todd G.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Cooper, Bret
 ; APPLICANT: Glazebrook, Jane
 ; APPLICANT: Goff, Stephen A.
 ; APPLICANT: Katagiri, Fumiaki
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Provart, Nicholas
 ; APPLICANT: Ricke, Darrell
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
 ; FILE REFERENCE: 60111-NP
 ; CURRENT APPLICATION NUMBER: US/10/260,238
 ; CURRENT FILING DATE: 2002-09-26
 ; PRIOR APPLICATION NUMBER: US 60/325,448
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: US 60/325,277
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: US 60/370,620
 ; PRIOR FILING DATE: 2002-04-04
 ; NUMBER OF SEQ ID NOS: 6077
 ; SEQ ID NO 1261
 ; LENGTH: 1325
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; US-10-260-238-1261

Query Match 1.5%; Score 18; DB 15; Length 1325;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 CTCCTGAAGATCCTCCC 219
 DB 639 CTCCTGAAGATCCTCCC 656

RESULT 73
 US-09-997-701-4
 ; Sequence 4, Application US/09997701
 ; Patent No. US20020107180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Gorgone, Gina A.
 ; APPLICANT: Baughn, Mariah R.
 ; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
 ; FILE REFERENCE: PF-0631 US
 ; CURRENT APPLICATION NUMBER: US/09/997,701
 ; CURRENT FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/470,946
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-22
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 4
 ; LENGTH: 1438
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: 2705267
 ; US-09-997-701-4

Query Match 1.5%; Score 18; DB 9; Length 1438;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGATGAC 1156
 DB 855 TACATCAGCCTGATGAC 872

```

; PRIOR APPLICATION NUMBER: 09/311832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/380142
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/644848
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 09/747259
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/816744
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/854208
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/854280
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/874503
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 09/869599
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/908,827
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US99/10733
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30720
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: PCT/US00/05601
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/34956
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: PCT/US01/17800
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/21066
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 80
; SEQ ID NO 8
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-931-836-8

Query Match      1.5%; Score 18; DB 10; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1139 TACATCAGCCTGAATGAC 1156
DB      855 TACATCAGCCTGAATGAC 872

RESULT 75
US-10-206-915-441
; Sequence 441, Application US/10206915
; Publication No. US200409221A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C513
; CURRENT APPLICATION NUMBER: US/10/206,915
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-199-670-441

```

```

Query Match 1.5%: Score 18; DB 12; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

```

```

RESULT 76
US-10-199-670-441
; Sequence 441, Application US/10199670
; Publication No. US20040033560A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C401
; CURRENT APPLICATION NUMBER: US/10/199,670

```

```

; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-199-670-441

```

```

Query Match 1.5%: Score 18; DB 12; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

```

```

RESULT 77
US-10-201-858-441
; Sequence 441, Application US/10201858
; Publication No. US20040038337A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C464
; CURRENT APPLICATION NUMBER: US/10/201,858
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21

```

```

; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-201-858-441

Query Match 1.5%; Score 18; DB 12; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 79
US-10-208-024-441
; Sequence 441, Application US/10208024
; Publication No. US20040048335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P34301C538
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-208-024-441

Query Match 1.5%; Score 18; DB 12; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 80
US-10-201-853-441
; Sequence 441, Application US/10201853
; Publication No. US20040053358A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
```

```

; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C465
; CURRENT APPLICATION NUMBER: US/10/201,853
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-201-853-441

Query Match 1.5%; Score 18; DB 12; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 81
US-10-036-342-8
; Sequence 8, Application US/10036342
; Publication No. US20020090681A1
; GENERAL INFORMATION:
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3030R1C5
; CURRENT APPLICATION NUMBER: US/10/036,342
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: PCT/US99/10733
```



```

; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30720
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: PCT/US00/05601
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/34956
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/17800
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/21066
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 80
; SEQ ID NO 8
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-036-342-8

Query Match      1.5%; Score 18; DB 13; Length 1579;
Best Local Similarity 100.0%; Fred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 82
US-10-052-586-441
; Sequence 441, Application US/10052586
; Publication No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
```

PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-16
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088940

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089908

Query Match 1.5%; Score 18; DB 13; Length 1579;

Best Local Similarity 100.0%; Pred.No. 94; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156

Db 855 TACATCAGCCTGAATGAC 872

RESULT 83

US-10-036-041-8
Sequence 8, Application US/10036041
Publication No. US20020192751A1
GENERAL INFORMATION:
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Stewart, Timothy A.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3030R1C8
CURRENT APPLICATION NUMBER: US/10/036,041
CURRENT FILING DATE: 2001-12-26
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/112514
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140

;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115552
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/116843
;; PRIOR FILING DATE: 1999-01-22
;; PRIOR APPLICATION NUMBER: 60/125774
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/125778
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/125826
;; PRIOR FILING DATE: 1999-03-24
;; PRIOR APPLICATION NUMBER: 60/127035
;; PRIOR FILING DATE: 1999-03-31
;; PRIOR APPLICATION NUMBER: 60/127706
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/129122
;; PRIOR FILING DATE: 1999-04-13
;; PRIOR APPLICATION NUMBER: 60/130359
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131272
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/132371
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/132379
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/132383
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/135750
;; PRIOR FILING DATE: 1999-05-25
;; PRIOR APPLICATION NUMBER: 60/138166
;; PRIOR FILING DATE: 1999-06-08
;; PRIOR APPLICATION NUMBER: 60/144791
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/146970
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/162506
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: 09/311832
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 09/380142
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/644848
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 09/747259
;; PRIOR FILING DATE: 2000-12-20
;; PRIOR APPLICATION NUMBER: 09/816744
;; PRIOR FILING DATE: 2001-03-22
;; PRIOR APPLICATION NUMBER: 09/854208
;; PRIOR FILING DATE: 2001-05-10
;; PRIOR APPLICATION NUMBER: 09/854280
;; PRIOR FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: 09/874503
;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: 09/869599
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: 09/908,827
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: PCT/US99/10733
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: PCT/US99/28551
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30720
;; PRIOR FILING DATE: 1999-12-22
;; PRIOR APPLICATION NUMBER: PCT/US00/05601
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: PCT/US00/14042
;; PRIOR FILING DATE: 2000-05-22

;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/23522
;; PRIOR FILING DATE: 2000-08-23
;; PRIOR APPLICATION NUMBER: PCT/US00/23328
;; PRIOR FILING DATE: 2000-08-24
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; PRIOR APPLICATION NUMBER: PCT/US00/34956
;; PRIOR FILING DATE: 2000-12-20
;; PRIOR APPLICATION NUMBER: PCT/US01/06520
;; PRIOR FILING DATE: 2001-02-28
;; PRIOR APPLICATION NUMBER: PCT/US01/17800
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: PCT/US01/19692
;; PRIOR FILING DATE: 2001-06-20
;; PRIOR APPLICATION NUMBER: PCT/US01/21066
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: PCT/US01/21735
;; PRIOR FILING DATE: 2001-07-09
;; NUMBER OF SEQ ID NOS: 80
;; SEQ ID NO 8
;; LENGTH: 1579
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-036-041-8

Query Match 1.5%; Score 18; DB 13; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCTGAATGAC 1156
|||||
Db 855 TACATCAGCTGAATGAC 872

RESULT 84

US-10-035-855-8
; Sequence 8, Application US/10035855
; Publication No. US2003008348A1
; GENERAL INFORMATION:
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3030R1C4
; CURRENT APPLICATION NUMBER: US/10/035,855
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/112514
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113300
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113430
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/114140
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115552
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/116843

```
/ PRIOR FILING DATE: 1999-01-22
/ PRIOR APPLICATION NUMBER: 60/125774
/ PRIOR FILING DATE: 1999-03-23
/ PRIOR APPLICATION NUMBER: 60/125778
/ PRIOR FILING DATE: 1999-03-23
/ PRIOR APPLICATION NUMBER: 60/125826
/ PRIOR FILING DATE: 1999-03-24
/ PRIOR APPLICATION NUMBER: 60/127035
/ PRIOR FILING DATE: 1999-03-31
/ PRIOR APPLICATION NUMBER: 60/127706
/ PRIOR FILING DATE: 1999-04-05
/ PRIOR APPLICATION NUMBER: 60/129122
/ PRIOR FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: 60/130359
/ PRIOR FILING DATE: 1999-04-21
/ PRIOR APPLICATION NUMBER: 60/131270
/ PRIOR FILING DATE: 1999-04-27
/ PRIOR APPLICATION NUMBER: 60/131272
/ PRIOR FILING DATE: 1999-04-27
/ PRIOR APPLICATION NUMBER: 60/131291
/ PRIOR FILING DATE: 1999-04-27
/ PRIOR APPLICATION NUMBER: 60/132371
/ PRIOR FILING DATE: 1999-05-04
/ PRIOR APPLICATION NUMBER: 60/132379
/ PRIOR FILING DATE: 1999-05-04
/ PRIOR APPLICATION NUMBER: 60/132383
/ PRIOR FILING DATE: 1999-05-04
/ PRIOR APPLICATION NUMBER: 60/135750
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: 60/138166
/ PRIOR FILING DATE: 1999-06-08
/ PRIOR APPLICATION NUMBER: 60/144791
/ PRIOR FILING DATE: 1999-07-20
/ PRIOR APPLICATION NUMBER: 60/146970
/ PRIOR FILING DATE: 1999-08-03
/ PRIOR APPLICATION NUMBER: 60/162506
/ PRIOR FILING DATE: 1999-10-29
/ PRIOR APPLICATION NUMBER: 09/311832
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: 09/380142
/ PRIOR FILING DATE: 1999-08-25
/ PRIOR APPLICATION NUMBER: 09/644848
/ PRIOR FILING DATE: 2000-08-22
/ PRIOR APPLICATION NUMBER: 09/747259
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 09/816744
/ PRIOR FILING DATE: 2001-03-22
/ PRIOR APPLICATION NUMBER: 09/854208
/ PRIOR FILING DATE: 2001-05-10
/ PRIOR APPLICATION NUMBER: 09/854280
/ PRIOR FILING DATE: 2001-05-10
/ PRIOR APPLICATION NUMBER: 09/874503
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: 09/869599
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: 09/908,827
/ PRIOR FILING DATE: 2001-07-18
/ PRIOR APPLICATION NUMBER: PCT/US99/10733
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: PCT/US99/28551
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/30720
/ PRIOR FILING DATE: 1999-12-22
/ PRIOR APPLICATION NUMBER: PCT/US00/05601
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: PCT/US00/05841
/ PRIOR FILING DATE: 2000-03-02
/ PRIOR APPLICATION NUMBER: PCT/US00/14042
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: PCT/US00/15264
/ PRIOR FILING DATE: 2000-06-02
/ PRIOR APPLICATION NUMBER: PCT/US00/23522
/ PRIOR FILING DATE: 2000-08-23
```

```
/ PRIOR APPLICATION NUMBER: PCT/US00/23328
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: PCT/US00/32678
/ PRIOR FILING DATE: 2000-12-01
/ PRIOR APPLICATION NUMBER: PCT/US00/34956
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: PCT/US01/06520
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: PCT/US01/17800
/ PRIOR FILING DATE: 2001-06-01
/ PRIOR APPLICATION NUMBER: PCT/US01/19692
/ PRIOR FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: PCT/US01/21066
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: PCT/US01/21735
/ PRIOR FILING DATE: 2001-07-09
/ NUMBER OF SEQ ID NOS: 80
/ SEQ ID NO 8
/ LENGTH: 1579
/ TYPE: DNA
/ ORGANISM: Homo Sapien
US-10-035-855-8
```

Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872
```

```
RESULT 85
US-10-174-590-441
; Sequence 441, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-590-441
```

Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872
```

```
RESULT 86
US-10-176-758-441
; Sequence 441, Application US/10176758
```

```

; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C04
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-758-441

Query Match      1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 87
US-10-175-737-441
; Sequence 441, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-737-441

Query Match      1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

```

```

RESULT 88
US-10-173-706-441
; Sequence 441, Application US/10173706
; Publication No. US2003002293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-441

Query Match      1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 89
US-10-175-738-441
; Sequence 441, Application US/10175738
; Publication No. US2003002294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-738-441

Query Match      1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156

```

Db 855 TACATCAGCCTGAATGAC 872

RESULT 90

US-10-175-752-441
; Sequence 441, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175, 752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-752-441

Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 91

US-10-176-482-441
; Sequence 441, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176, 482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-482-441

Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 92

US-10-176-757-441
; Sequence 441, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176, 757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-757-441

Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 93

US-10-176-913-441
; Sequence 441, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176, 913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-913-441

```
Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 94
US-10-180-552-441
; Sequence 441, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-552-441

Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 95
US-10-180-557-441
; Sequence 441, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180,557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-557-441
```

```
; ORGANISM: Homo Sapien
US-10-180-557-441

Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 96
US-10-173-700-441
; Sequence 441, Application US/10173700
; Publication No. US20030027262A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C14
; CURRENT APPLICATION NUMBER: US/10/173,700
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-700-441

Query Match 1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 97
US-10-174-572-441
; Sequence 441, Application US/10174572
; Publication No. US20030027263A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C40
; CURRENT APPLICATION NUMBER: US/10/174,572
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-572-441
```

```
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-572-441

Query Match
  1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 98
US-10-174-579-441
; Sequence 441, Application US/10174579
; Publication No. US20030027264A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C31
; CURRENT APPLICATION NUMBER: US/10/174,579
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-579-441

Query Match
  1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 99
US-10-174-582-441
; Sequence 441, Application US/10174582
; Publication No. US20030027265A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C36
; CURRENT APPLICATION NUMBER: US/10/174,582
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-582-441

Query Match
  1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

RESULT 100
US-10-174-588-441
; Sequence 441, Application US/10174588
; Publication No. US20030027266A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C28
; CURRENT APPLICATION NUMBER: US/10/174,588
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 441
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-588-441

Query Match
  1.5%; Score 18; DB 14; Length 1579;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1139 TACATCAGCCTGAATGAC 1156
Db 855 TACATCAGCCTGAATGAC 872

Search completed: March 25, 2004, 10:03:00
Job time : 465 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 05:52:16 ; Search time 3491 Seconds
(without alignments)
10119.438 Million cell updates/sec

Title: US-09-939-853A-74
Perfect score: 1183
Sequence: 1 agctagagcctccaggagcc.....tctcttgatgatgactag 1183

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_pig:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	735	62.1	1201	9	AL541041
2	647	54.7	1002	12	BQ052308
3	578	48.9	986	12	BQ054265
4	568	48.0	1020	12	BQ054281

554	46.8	1069	12	BQ052468	AGENCOURT
459	38.8	516	13	EX383606	EX383606
426	36.0	778	12	BG178487	BG178487
405	34.2	566	12	BG284179	BG284179
390	33.0	878	12	BQ053486	BQ053486
300	25.4	597	9	AL844311	AL844311
276	23.3	614	9	AL844307	AL844307
255	21.6	794	12	BG677567	BG677567
205	17.3	560	13	BQ944126	BQ944126
14	16.9	606	9	AL844309	AL844309
116	9.8	642	9	AL844308	AL844308
106	9.0	611	9	AL844310	AL844310
103	8.7	506	28	B45150	B45150
80	6.8	619	9	AL844312	AL844312
68	5.7	546	28	AQ556467	AQ556467
57	4.8	149	9	AA309769	AA309769
47	4.0	553	28	AQ556478	AQ556478
43	3.6	660	10	B8635615	B8635615
43	3.6	660	13	BY742155	BY742155
43	3.6	926	11	AK020837	AK020837
43	3.6	2637	11	AK088672	AK088672
43	3.6	2974	11	AK030877	AK030877
42	3.6	322	10	AW437301	AW437301
42	3.6	569	14	CB426333	CB426333
34	2.9	377	9	AA959151	AA959151
29	2.5	389	10	BE015229	BE015229
29	2.5	603	13	BQ553005	BQ553005
28	2.4	627	10	BB619854	BB619854
27	2.3	757	29	CC527971	CC527971
23	1.9	541	12	B1898999	B1898999
22	1.9	764	12	BG388843	BG388843
22	1.9	791	14	CB994934	CB994934
22	1.9	875	12	BG282272	BG282272
22	1.9	934	12	B1549288	B1549288
22	1.9	1109	12	BM926459	BM926459
22	1.9	2517	29	AY418068	AY418068
22	1.9	2517	29	AY418069	AY418069
21	1.8	320	14	CB657851	CB657851
21	1.8	332	10	BE030537	BE030537
21	1.8	363	9	AA365070	AA365070
21	1.8	408	29	CG869927	CG869927
21	1.8	454	14	CF540846	CF540846
21	1.8	561	28	AQ372880	AQ372880
21	1.8	653	28	AZ716247	AZ716247
21	1.8	661	29	CE297560	CE297560
21	1.8	668	10	BB629807	BB629807
21	1.8	821	12	B1752759	B1752759
21	1.8	840	28	BZ02783	BZ02783
21	1.8	2296	11	AK009075	AK009075
21	1.8	2311	11	AK036802	AK036802
21	1.8	163	14	H75025	H75025
21	1.7	282	10	BB563180	BB563180
21	1.7	341	9	AA589496	AA589496
21	1.7	347	14	CB705122	CB705122
21	1.7	376	9	AI585555	AI585555
21	1.7	381	14	CB810150	CB810150
21	1.7	388	9	AA510624	AA510624
21	1.7	411	14	CB802633	CB802633
21	1.7	425	9	AA271813	AA271813
21	1.7	427	10	BF533481	BF533481
21	1.7	439	10	BE015226	BE015226
21	1.7	449	14	CF391835	CF391835
21	1.7	455	13	BQ198815	BQ198815
21	1.7	462	10	AW783749	AW783749
21	1.7	462	12	BG555321	BG555321
21	1.7	462	14	CA728751	CA728751
21	1.7	477	10	BE2688290	BE2688290
21	1.7	478	9	AA547055	AA547055
21	1.7	485	14	CA545265	CA545265
21	1.7	507	9	AA051232	AA051232
21	1.7	525	9	AU127488	AU127488
21	1.7	526	9	AA818383	AA818383
21	1.7	549	14	CD892975	CD892975

78	20	1.7	552	10	AM647304	EST325847	151	19	1.6	240	9	AJ481865	AJ481865
79	20	1.7	553	12	BJ632593	BJ632593	152	19	1.6	240	9	AV913590	AV913590
80	20	1.7	554	14	CB586549	AMGNNUC:N	C 153	19	1.6	246	14	D81734	D81734
81	20	1.7	602	14	CB115016	K-EST0158	C 154	19	1.6	246	9	AI175468	AI175468
82	20	1.7	619	10	BE694262	602082893	155	19	1.6	248	14	CA628885	CA628885
83	20	1.7	619	28	AQ420273	RPCI-11-1	156	19	1.6	256	14	CA726301	CA726301
84	20	1.7	640	13	BQ553006	H019E02-	157	19	1.6	257	14	CA735447	CA735447
85	20	1.7	653	13	BY746881	BY746881	158	19	1.6	259	9	AL829315	AL829315
86	20	1.7	654	13	BW311942	BW311942	159	19	1.6	259	14	CA726150	CA726150
87	20	1.7	655	10	BE261614	601149190	160	19	1.6	260	14	CA737840	CA737840
88	20	1.7	670	9	AI510095	mj43C04_Y	161	19	1.6	260	14	CA737934	CA737934
89	20	1.7	689	13	BM305486	BM305486	162	19	1.6	260	14	CA739148	CA739148
90	20	1.7	696	28	BZ828172	PUGR688TD	163	19	1.6	260	14	CA739591	CA739591
91	20	1.7	710	9	AV968086	AV968086	C 164	19	1.6	262	14	CA735254	CA735254
92	20	1.7	710	10	BF797024	602258062	165	19	1.6	263	14	CA739332	CA739332
93	20	1.7	719	13	BX642175	DKFZP686B	166	19	1.6	273	9	AV171559	AV171559
94	20	1.7	736	12	B1153749	602871407	167	19	1.6	278	9	AV171530	AV171530
95	20	1.7	737	9	AV705440	AV705440	C 168	19	1.6	280	10	BI108007	BI108007
96	20	1.7	746	28	BE282167	PUGR688TB	169	19	1.6	287	12	BI433272	BI433272
97	20	1.7	756	10	BE256881	601112037	170	19	1.6	287	13	CA021764	CA021764
98	20	1.7	763	12	BE873348	602794320	171	19	1.6	300	9	AJ481862	AJ481862
99	20	1.7	764	12	BI084247	602869751	172	19	1.6	300	9	AJ481866	AJ481866
100	20	1.7	773	28	BE774469	mcv95804_Y	173	19	1.6	307	12	BJ552097	BJ552097
101	20	1.7	775	13	BM176012	BM176012	C 174	19	1.6	319	9	AA632853	AA632853
102	20	1.7	783	10	BE396636	601289120	175	19	1.6	319	12	BM134399	BM134399
103	20	1.7	783	12	BI820532	603036164	176	19	1.6	320	14	CA734184	CA734184
104	20	1.7	799	12	BI225407	602950162	C 177	19	1.6	327	9	AI736882	AI736882
105	20	1.7	800	12	BI833016	603087325	C 178	19	1.6	331	9	AA144868	AA144868
106	20	1.7	803	9	AUI25413	AUI25413	179	19	1.6	335	10	AW259857	AW259857
107	20	1.7	803	10	BF665822	602124265	180	19	1.6	341	10	AW647528	AW647528
108	20	1.7	814	12	BM947197	UI-W-RHOP	181	19	1.6	347	28	BZ951352	BZ951352
109	20	1.7	819	28	BE259304	WACAP46TR	182	19	1.6	347	10	BE406046	BE406046
110	20	1.7	824	10	BF125084	60181531	183	19	1.6	349	10	BE406046	BE406046
111	20	1.7	826	13	BU524873	BU524873	184	19	1.6	352	13	CA029217	CA029217
112	20	1.7	830	10	BE742805	601574424	185	19	1.6	355	10	BE518158	BE518158
113	20	1.7	830	12	BE756739	602710210	186	19	1.6	360	9	AJ470889	AJ470889
114	20	1.7	853	10	BF696156	602124669	187	19	1.6	360	9	AJ481863	AJ481863
115	20	1.7	854	12	BE756534	602715677	188	19	1.6	360	9	AJ481867	AJ481867
116	20	1.7	859	10	BE141756	601787621	189	19	1.6	361	9	AJ481869	AJ481869
117	20	1.7	867	12	BI104925	602892722	190	19	1.6	361	13	BQ579495	BQ579495
118	20	1.7	868	13	BE406526	BE406526	C 191	19	1.6	361	14	DE1508	DE1508
119	20	1.7	873	29	CNS02TKO	AL213273 Tetraodon	192	19	1.6	362	9	AJ601702	AJ601702
120	20	1.7	912	12	BE686106	602638679	193	19	1.6	362	10	AW227209	AW227209
121	20	1.7	916	13	BQ947590	AGENCOURT	194	19	1.6	365	13	BQ761361	BQ761361
122	20	1.7	925	10	BE300734	602031917	195	19	1.6	366	13	BY021522	BY021522
123	20	1.7	933	14	CB205297	AGENCOURT	C 196	19	1.6	367	9	AI112486	AI112486
124	20	1.7	937	29	CG278194	CG278194	197	19	1.6	367	9	AL819200	AL819200
125	20	1.7	940	13	BU902817	BU902817	198	19	1.6	367	14	CA718844	CA718844
126	20	1.7	941	12	BI413430	602990326	199	19	1.6	372	14	CA614298	CA614298
127	20	1.7	949	12	BE341384	602463865	200	19	1.6	375	10	BF291388	BF291388
128	20	1.7	956	13	BQ959172	AGENCOURT	201	19	1.6	376	9	AI862103	AI862103
129	20	1.7	960	12	BE6968201	602835685	202	19	1.6	377	10	AW549208	AW549208
130	20	1.7	964	12	BE404521	602420656	203	19	1.6	378	13	BY020881	BY020881
131	20	1.7	966	12	BE919471	AGENCOURT	204	19	1.6	381	9	AL830425	AL830425
132	20	1.7	998	10	BF129162	601811431	205	19	1.6	382	12	BI961171	BI961171
133	20	1.7	1023	12	BE760066	602733305	C 206	19	1.6	384	14	DE1119	DE1119
134	20	1.7	1028	12	BE917288	AGENCOURT	207	19	1.6	386	9	AI956576	AI956576
135	20	1.7	1052	29	CNS04243	Tetraodon	208	19	1.6	386	9	BI846189	BI846189
136	20	1.7	1055	12	BE918144	AGENCOURT	209	19	1.6	389	12	BJ301731	BJ301731
137	20	1.7	1071	10	BE511173	601235837	210	19	1.6	389	14	W18283	W18283
138	20	1.7	1113	12	BE477949	AGENCOURT	211	19	1.6	395	9	AL828959	AL828959
139	20	1.7	1192	12	BE545575	AGENCOURT	212	19	1.6	396	12	BU232190	BU232190
140	20	1.7	1201	9	AL557199	AL557199	213	19	1.6	399	13	BU985479	BU985479
141	19	1.6	70	9	AA588675	nm98d05_S	214	19	1.6	400	12	BJ463398	BJ463398
142	19	1.6	95	13	BO760009	EBP107_S	215	19	1.6	405	13	BY006369	BY006369
143	19	1.6	183	13	CA028377	HZ61M20r	216	19	1.6	406	10	AW259327	AW259327
144	19	1.6	197	14	CB059853	4009706_B	217	19	1.6	407	10	AW394806	AW394806
145	19	1.6	197	14	CB060475	4011101_B	C 218	19	1.6	407	10	BE107097	BE107097
146	19	1.6	202	14	CB105648	K-EST0132	219	19	1.6	409	10	BB846348	BB846348
147	19	1.6	208	12	BE464732	BJ464732	220	19	1.6	409	12	BU224315	BU224315
148	19	1.6	215	14	CB105463	CB105463	C 221	19	1.6	409	12	BJ231789	BJ231789
149	19	1.6	219	14	CB105464	K-EST0128	222	19	1.6	410	12	BJ208301	BJ208301
150	19	1.6	240	9	AJ481864	AJ481864	223	19	1.6	413	12	BMC99714	BMC99714

```
224 19 1.6 414 9 AL638913
225 19 1.6 416 9 AL829273
226 19 1.6 418 9 AJ433205
227 19 1.6 418 10 BB848131
228 19 1.6 419 9 AJ552134
229 19 1.6 419 14 CA641064
230 19 1.6 420 9 AJ460007
231 19 1.6 420 9 AJ460759
232 19 1.6 420 9 AJ470888
233 19 1.6 420 9 AJ481861
234 19 1.6 420 9 AJ481868
235 19 1.6 420 9 AJ481870
236 19 1.6 422 13 BY006564
237 19 1.6 424 9 AB000302
238 19 1.6 425 10 BE425808
239 19 1.6 427 14 CA713924
240 19 1.6 430 10 BF202744
241 19 1.6 430 10 BF231354
242 19 1.6 431 14 CA785016
243 19 1.6 432 9 A610201
244 19 1.6 433 13 BU979892
245 19 1.6 433 14 CA713656
246 19 1.6 435 14 CA720633
247 19 1.6 435 14 CA729273
248 19 1.6 436 14 CA648003
249 19 1.6 436 14 CA710059
250 19 1.6 437 10 BB847450
251 19 1.6 438 13 CA025642
252 19 1.6 438 14 CA717898
253 19 1.6 439 14 CA711217
254 19 1.6 441 10 BB845915
255 19 1.6 441 14 CA714356
256 19 1.6 442 9 A1125308
257 19 1.6 442 12 BG881364
258 19 1.6 444 13 BU981322
259 19 1.6 446 14 CA719892
260 19 1.6 446 28 AQ807361
261 19 1.6 447 9 A1171082
262 19 1.6 447 14 CA670432
263 19 1.6 448 9 A1506844
264 19 1.6 448 14 CA668084
265 19 1.6 448 14 CA723147
266 19 1.6 450 14 CB787692
267 19 1.6 451 12 BM447177
268 19 1.6 453 9 A1139136
269 19 1.6 455 13 BV650337
270 19 1.6 456 12 BU227079
271 19 1.6 457 14 CA612455
272 19 1.6 459 12 BU258118
273 19 1.6 460 9 AU205635
274 19 1.6 462 9 AA987061
275 19 1.6 462 14 CD465234
276 19 1.6 464 10 AW258820
277 19 1.6 465 10 B5490693
278 19 1.6 467 28 AZ856474
279 19 1.6 468 6 AL810412
280 19 1.6 468 9 A1125365
281 19 1.6 468 13 BO756665
282 19 1.6 469 13 BU989211
283 19 1.6 469 14 CB105644
284 19 1.6 469 29 CE345261
285 19 1.6 470 9 AL778486
286 19 1.6 470 9 AL830666
287 19 1.6 470 14 CA668823
288 19 1.6 472 12 BM373928
289 19 1.6 472 14 CD465199
290 19 1.6 473 14 CA608356
291 19 1.6 473 14 CD888122
292 19 1.6 475 10 BF483983
293 19 1.6 475 14 CA720764
294 19 1.6 477 10 AW526152
295 19 1.6 477 13 BU979872
296 19 1.6 479 14 CD886596

AL638913 AL638913
AL829273 AL829273
AJ433205 AJ433205
BB848131 BB848131
AJ552134 mm92hl0.y
CA641064 wreln.pk0
AJ460007 AJ460007
AJ460759 AJ460759
AJ470888 AJ470888
AJ481861 AJ481861
AJ481868 AJ481868
AJ481870 AJ481870
BY006564 BY006564
AB000302 EST189799
BE425808 WHE0337.E
CA713924 wkm2n.pk0
BF202744 WHE1784.A
BP231354 253724.BA
CA785016 sat70c08.
A610201 AJ610201
BU979892 HA17L08r
CA713656 wdk3c.pk0
CA720633 wkm2n.pk0
CA729273 wdlc.pk0
CA648003 wreln.pk0
CA710059 wdk2c.pk0
BB847450 BB847450
CA025642 H25L21r
CA717898 wdk5c.pk0
CA711217 wdk2c.pk0
BB845915 BB845915
CA714356 wdk3c.pk0
A1125308 qd88c01.x
BG881364 sae81b05.
BU981322 HA23E13r
CA719892 wkm2n.pk0
AQ807361 HS 3177.A
A1171082 EST217030
CA670432 wlsul.pk0
A1506844 A1506844
CA668084 wlsul.pk0
CA723147 wdlf.pk0
CB787692 AMGNNUC:N
BM447177 DSA006H09
A1139136 qci9b11.x
BV650337 BV650337
BU227079 BU227079
CA612455 wri.pk013
BU258118 BU258118
AU205635 AU205635
AA987061 uc74g09.x
CD465234 LeukoN1.2
AW258820 um74b02.Y
B5490693 WHE0369.E
AZ856474 2M0160H24
AL810412 Triticum
A1125365 qd88e07.x
BO756665 EBem09.SQ
BU989211 HE21BJ5r
CB105644 K.FST0131
CE345261 C1gr-9ss-
AL778486 AL778486
AL830666 AL830666
CA668823 wlsul.pk0
BM373928 EBma03.SQ
CD465199 LeukoN2
CA608356 wri.pk006
CD888122 G118.107C
BF483983 WHE2306.H
CA720764 wkm2n.pk0
AW526152 UI-R-B01-
BU979872 HA17K10r
CD886596 G118.102L
```

ALIGNMENTS

RESULT 1

```
AL541041 1201 bp mRNA linear EST 12-MAY-2003
LOCUS AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YK23
DEFINITION S-PRIME, mRNA sequence.
ACCESSION AL541041
VERSION AL541041.2 GI:30544829
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12871733.
Contact: Genoscope, Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9825.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE005AF12Q01&cluster=9825.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005AF12Q01.
```

FEATURES

```
1..1201
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YK23"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
```

ORIGIN

```
Query Match 62.1%; Score 735; DB 9; Length 1201;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 835; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TAGAGCTCCAAAGACCCACCGCTGTCTCTCTGTGACAGAGCTCAAGGGCCCTGGSCCT 63
Db 65 TAGAGCTCCAAAGACCCACCGCTGTCTCTCTGTGACAGAGCTCAAGGGCCCTGGSCCT 124
Qy 64 TCCCTCCCTGGCTGGCTGTCTGTGGAGGGTTCCTCCAGATCCCTAGGAGCAT 123
Db 125 TCCCTCCCTGGCTGGCTGTCTGTGGAGGGTTCCTCCAGATCCCTAGGAGCAT 184
Qy 124 GGGGACGCTATCATCCCTGGGTGTACAACTGCTACTGCACAGAGATGCTAGCTACC 183
Db 185 GGGGACGCTATCATCCCTGGGTGTACAACTGCTACTGCACAGAGATGCTAGCTACC 244
Qy 184 CAAACCAACACCTAGCTCTCCCTGAGAGATCCCTCCAGGCTGAGAGATGCTGGGTGCC 243
Db 245 CAAACCAACACCTAGCTCTCCCTGAGAGATCCCTCCAGGCTGAGAGATGCTGGGTGCC 304
```

QY 244 TAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTTACCTGTCCAGCC 303
 Db 305 TAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTTACCTGTCCAGCC 364
 QY 304 AGAGCATGGCTCTCAGCAGAGCTGTCTTCCCAAGCCCTTTCAGTACAAACCAATTTCCCTC 363
 Db 365 AGAGCATGGCTCTCAGCAGAGCTGTCTTCCCAAGCCCTTTCAGTACAAACCAATTTCCCTC 424
 QY 364 GATGATGCTCTCTGAGTCTCTGAGGACCAATGGGAAGTCTGCCAGCAGAGAAA 423
 Db 425 GATGATGCTCTCTGAGTCTCTGAGGACCAATGGGAAGTCTGCCAGCAGAGAAA 484
 QY 424 ATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTCAGAGCCAGGACCTGTGACCATGGA 483
 Db 485 ATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTCAGAGCCAGGACCTGTGACCATGGA 544
 QY 484 AGCAGAGAGAAGCAAGCCACAGCCGTCGCTGGGCAAGTTTCCCGCAGGTGGCCCGC 543
 Db 545 AGCAGAGAGAAGCAAGCCACAGCCGTCGCTGGGCAAGTTTCCCGCAGGTGGCCCGC 604
 QY 544 CGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGATGGAGCTGGT 603
 Db 605 CGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGATGGAGCTGGT 664
 QY 604 GACGCTGTCTCTGAACTCTCAGGAGAGAGATATAACATCTCCAGCGTCCAGTGGCAA 663
 Db 665 GACGCTGTCTCTGAACTCTCAGGAGAGAGATATAACATCTCCAGCGTCCAGTGGCAA 724
 QY 664 AGTCTCCATGGTGTCTGTATGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 723
 Db 725 AGTCTCCATGGTGTCTGTATGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 784
 QY 724 GTTACTCTGAGTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 783
 Db 785 GTTACTCTGAGTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 844
 QY 784 TTTACTCTGAGTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 Db 845 TTTACTCTGAGTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 901

RESULT 2
 BQ052308 1002 bp mRNA linear EST 29-MAR-2002
 LOCUS AGENCOURT_6868571 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933542
 5', mRNA sequence.

ACCESSION BQ052308
 VERSION BQ052308.1 GI:19811648

KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1002)
 NIH-MGC <http://mgi.nhl.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

AUTHORS Contact: Robert Strausberg, Ph.D.
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 Email: cgabbs-r@mail.nih.gov
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLCM2118 row: d column: 23
 High quality sequence stop: 670.
 Location/Qualifiers

FEATURES
 source

1..1002
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

/clone="IMAGE:5933542"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_106"
 /note="Organ: blood; Vector: pOTB7; Site: 1; XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Query Match 54.7%; Score 647; DB 12; Length 1002;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 747; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 333 CCAAGCCCTTTGATGACAAACCAATTTCCCTCGATGATGTGCTTCTGAGTCTCTGCTGAG 392
 Db 31 CCAAGCCCTTTGATGACAAACCAATTTCCCTCGATGATGTGCTTCTGAGTCTCTGCTGAG 90
 QY 393 GAACAATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAATTCCT 452
 Db 91 GAACAATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAATTCCT 150
 QY 453 CTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCAAGGCCACAGCCGTGG 512
 Db 151 CTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCAAGGCCACAGCCGTGG 210
 QY 513 CCCTGGGCACTTCCCGGCAAGTGGCCCGCCGAGCTGCTGAGACTCGGGAGGCCAT 572
 Db 211 CCCTGGGCACTTCCCGGCAAGTGGCCCGCCGAGCTGCTGAGACTCGGGAGGCCAT 270
 QY 573 TGACCATCTCTCTGAGGATGGAGACTGTGGAGCGTGTCTGAAGTCTCAGGCAGAG 632
 Db 271 TGACCATCTCTCTGAGGATGGAGACTGTGGAGCGTGTCTGAAGTCTCAGGCAGAG 330
 QY 633 AGTATAACATCCCCAGCGTCCAGTGGGCAAAAGTCTCCCATGGGTGCTGTATGAGGGCC 692
 Db 331 AGTATAACATCCCCAGCGTCCAGTGGGCAAAAGTCTCCCATGGGTGCTGTATGAGGGCC 390
 QY 693 TGAGCAGGAGAGACAGAGAGAACTGCTGTGTACTGGGAACCCCTGGAGGGCCCTTCC 752
 Db 391 TGAGCAGGAGAGAGACAGAGAGAACTGCTGTGTACTGGGAACCCCTGGAGGGCCCTTCC 450
 QY 753 TCATCCGGGAGAGCCAGACCCAGGAGAGGCTCTTACTCTCTGTCTGTCAGTCCGCTCAGCCGCC 812
 Db 451 TCATCCGGGAGAGCCAGACCCAGGAGAGGCTCTTACTCTCTGTCTGTCAGTCCGCTCAGCCGCC 510
 QY 813 CTGCATCTCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCCTTGACAAATGGCTGCTGT 872
 Db 511 CTGCATCTCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCCTTGACAAATGGCTGCTGT 570
 QY 873 ACATCTCACCAGCCCTCACCTTCCCTCACTCCAGGCGCTCTGTGGACCAATTAATCTCTGAGC 932
 Db 571 ACATCTCACCAGCCCTCACCTTCCCTCACTCCAGGCGCTCTGTGGACCAATTAATCTCTGAGC 630
 QY 933 TGGCGGATGACATCTGCTGCTACTCAAGAGGCCCTGTGTCTCTGACAGAGGGCTGGCCCGC 992
 Db 631 TGGCGGATGACATCTGCTGCTACTCAAGAGGCCCTGTGTCTCTGACAGAGGGCTGGCCCGC 690
 QY 993 TCCCTGGCAGGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCAACTGGAAG 1052
 Db 691 TCCCTGGCAGGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCAACTGGAAG 750
 QY 1053 AGCTGGACAGCTCCCTCCTCTCTTTCTGAA 1081
 Db 751 AGCTGGACAGCTCCCTCCTCTTTCTGAA 779

RESULT 3
 BQ054265

361	TTGAGCAGAGAGAGACAGAGCCACAGCCGTGGCCCTGGCAGTTTCCCGCAGGTGGCC	420
540	CGGCCGAGTGTGCGTGAAGATCTGGGGAGCCATTGACCATCGTCTCTGAGGATGGAGACT	599
421	CGGCCGAGTGTGCGTGAAGATCTGGGGAGCCATTGACCATCGTCTCTGAGGATGGAGACT	480
600	GGTGACCGTGTGTCGTGAAGTCTCAGGCAGAGAGTATAACATCCCAGCGTCCACGTGG	659
481	GGTGACCGTGTGTCGTGAAGTCTCAGGCAGAGAGTATAACATCCCAGCGTCCACGTGG	540
660	GCAAAGTCTCCATGGTGGCTGTATGAGGGCTGAGCAGGGAGAAAGCAGAGGAACCTGC	719
541	CCAAAGTCTCCATGGTGGCTGTATGAGGGCTGAGCAGGGAGAAAGCAGAGGAACCTGC	600
720	TCTGTATTACTGGAAACCTGGAGGGCC	748
601	TCTGTATTACTGGAAACCTGGAGGGCC	629

RESULT 4	ACCESSION	ORGANISM
BQ054281	VERSION	
LOCUS	KEYWORDS	
DEFINITION	SOURCE	

RESULT 4	1020 bp	linear	EST 29-MAR-2002
BQ054281		mRNA	
LOCUS			
AGENCOURT 6830234		Homo sapiens	cdna clone IMAGE:5936362
DEFINITION		5', mRNA sequence.	

```

FEATURES
source
1. .1020
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 106"
/notes="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

```

	Query Match	48.0%	Score 568	DB 12	Length 1020
	Best Local Similarity	99.6%	Pred. No. 1.5e-276		
	Matches 718	Conservative	0	Mismatches 3	Indels 0
Qy	191	ACACCTAGCCTCTCCCTGAAGATCTCCCGCGGTGAGAGATTCCTGGGTGTCCTAGGACC	250		
Db	5	ACACCTAGCCTCTCCCTGAAGATCTCCCGCGGTGAGAGATTCCTGGGTGTCCTAGGACC	64		

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L2CM2118 row: n column: 13
 High quality sequence stop: 681.

FEATURES
SOURCE

```

1. ...1069
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5933772"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 106"
/notes="organ: blood; vector: pOT87; Site 1: XhoI
EcoRI; cDNA made by oligo-dT priming. Direction
into EcoRI/XhoI sites using the following 5' a
GGCAGCAG(G). Library constructed by Ling Hong
laboratory of Gerald M. Rubin (University of
Berkeley) using ZAP-cDNA synthesis kit (Stratag
Superscript II RT (Life Technologies)). Note: t
NIH MGC library."

```

ORIGIN

Query Match	46.8%	Score 554;	DB 12;	Length 1069;
Best Local Similarity	99.8%	Pred. No. 1.9e-269;		
Matches 604;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps	0;			

Qy	310	TGGGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGGATCACAACCAATTTCCCTCGATGAT	369
Db	1	TGGGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGGATCACAACCAATTTCCCTCGATGAT	60
Qy	370	GTGCTTCTGAGTGTCTGTGCTGAGAAACAATGGGAAGTCTGCCAGCAGAGAAGAAATCTCT	429
Db	61	GTGCTTCTGAGTGTCTGTGCTGAGAAACAATGGGAAGTCTGCCAGCAGAGAAGAAATCTCT	120
Qy	430	GCCAAAGCCCAAGCTTGAGTTTCTCTGTCTCCAAAGCCAGGGACCTGTGACCATTGGAGCAGA	489
Db	121	GCCAAAGCCCAAGCTTGAGTTTCTCTGTCTCCAAAGCCAGGGACCTGTGACCATTGGAGCAGA	180
Qy	490	GAGAAGCAAGGCCACAGCCGTGSCCTTGGSCAGTTTCCGGCAGGTGGCCCGGCCGAGCT	549
Db	181	GAGAAGCAAGGCCACAGCCGTGSCCTTGGSCAGTTTCCGGCAGGTGGCCCGGCCGAGCT	240
Qy	550	GTGCTCTGAGACTCGGGGAGCCATTGACCATCTCTCTGAGGATGGAGACTGGTGGACGGT	609
Db	241	GTGCTCTGAGACTCGGGGAGCCATTGACCATCTCTCTGAGGATGGAGACTGGTGGACGGT	300
Qy	610	GCTGTCTGAAGTCTCAGGCAGAGATATAATCCCCAGCGTCCAATGGGGCAAAAGTCTC	669
Db	301	GCTGTCTGAAGTCTCAGGCAGAGATATAATCCCCAGCGTCCAATGGGGCAAAAGTCTC	360
Qy	670	CCATGGGTGGCTGTATGAGGGCCTGACGAGGAGAAACAGAGAACTGCTGTGTTTACC	729
Db	361	CCATGGGTGGCTGTATGAGGGCCTGACGAGGAGAAACAGAGAACTGCTGTGTTTACC	420
Qy	730	TGGGAACCTTGAAGGGGCCTTCTCATCCGGGAGAGCCAGACACGAGAGAGGCTCTTACT	789
Db	421	TGGGAACCTTGAAGGGGCCTTCTCATCCGGGAGAGCCAGACACGAGAGAGGCTCTTACT	480
Qy	790	TCTGTCAAGTCCGGCTCAGCCGCCCTGCATCCTGGGACCGGATCAGACACTACAGGATCCA	849
Db	481	TCTGTCAAGTCCGGCTCAGCCGCCCTGCATCCTGGGACCGGATCAGACACTACAGGATCCA	540
Qy	850	CTCCCTTGACAAATGGGTGGGTGTACATCTCACCGGCCCTCACCTTCCCTCACTCCAGGC	909
Db	541	CTCCCTTGACAAATGGGTGGGTGTACATCTCACCGGCCCTCACCTTCCCTCACTCCAGGC	600
Qy	910	CTTGG	914
Db	601	CTTGG	605

RESULT 6

251	QY		AAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTAA	CCTGTCCAGCCAGAGCAT	310
65	Db		AAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTAA	CCTGTCCAGCCAGAGCAT	124
311	QY		GGGTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAAT	TTCCCTCGATGATG	370
125	Db		GGGTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAAT	TTCCCTCGATGATG	184
371	QY		TGCTTCTGAGTCTCTGTGTGAGGAACAAATGGGAAGTCTGCCAGCAGAGAAAAT	CTCTCTG	430
185	Db		TGCTTCTGAGTCTCTGTGTGAGGAACAAATGGGAAGTCTGCCAGCAGAGAAAAT	CTCTCTG	244
431	QY		CCAAAGCCCAAAGCTTCAGTTCCTCTGTCCAAAGCCAGAGGACCTGTGACCAATGGGAAGCAGAG	490	
245	Db		CCAAAGCCCAAAGCTTCAGTTCCTCTGTCCAAAGCCAGAGGACCTGTGACCAATGGGAAGCAGAG	304	
491	QY		AGAAGCAAGGCCAACAGCCGTGGCCCTGGGCGAGTTTCCCGCAGGTGGCCCGCGCCGAGCTG	550	
305	Db		AGAAGCAAGGCCAACAGCCGTGGCCCTGGGCGAGTTTCCCGCAGGTGGCCCGCGCCGAGCTG	364	
551	QY		TGCTGTAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGACTGGTGGACGGTG	610	
365	Db		TGCTGTAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGACTGGTGGACGGTG	424	
611	QY		CTGCTTGAAGTCTCAGGCAGAGAGTATAACATCCCAGCGTCCACGTGGGCGAAAAGTCTCC	670	
425	Db		CTGCTTGAAGTCTCAGGCAGAGAGTATAACATCCCAGCGTCCACGTGGGCGAAAAGTCTCC	484	
671	QY		CATGGTGGCTGTATGAGGGCTCAGCAGGGAGAAAGCAGAGAACTGTGTGTTTACCT	730	
485	Db		CATGGTGGCTGTATGAGGGCTCAGCAGGGAGAAAGCAGAGAACTGTGTGTTTACCT	544	
731	QY		GGGAACCCCTGGAGGGGCCCTTCTCATCGGAGAGCCAGACACAGGAGAGGCTCTTACTCT	790	
545	Db		GGGAACCCCTGGAGGGGCCCTTCTCATCGGAGAGCCAGACACAGGAGAGGCTCTTACTCT	604	
791	QY		CTGTCAGTCCGGCTCAGCGGCCCTTGCAATCTTGGACCCGGATCAGACACTACAGGATCCAC	850	
605	Db		CTGTCAGTCCGGCTCAGCGGCCCTTGCAATCTTGGACCCGGATCAGACACTACAGGATCCAC	664	
851	QY		TGCCCTTGACAATGGCTGCTGTACATCTCACCAGCGGCTCACCTTCCCTCCTCAGGCG	910	
665	Db		TGCCCTTGACAATGGCTGCTGTACATCTCACCAGCGGCTCACCTTCCCTCCTCAGGCG	724	
911	QY		C	911	
725	Db		C	725	

RESULT 5
BQ052468
LOCUS
DEFINITION
1069 bp mRNA linear EST 29-MAR-2002
AGENCOURT 6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
5', mRNA sequence.
ACCESSION
BQ052468
VERSION
BQ052468.1 GI:19811808
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
1 (bases 1 to 1069)
TITLE
NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: grapbs-1@mail.nih.gov
Tissue procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

EX383606
LOCUS BX383606 616 bp mRNA linear EST 08-MAY-2003
DEFINITION BX383606 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 Homo sapiens cDNA clone CS0DJ013YK10 5-PRIME, mRNA sequence.
ACCESSION BX383606
VERSION BX383606.1 GI:30457152
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 616)
COMMENT Li.W.B., Gruber.C., Jesse,J. and Polayes.D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9825.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ013BF05QP1&cluster=9825.r. Contact :
 Feng liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DJ013BF05QP1.
FEATURES
 Location/Qualifiers
 1..616
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ013YK10"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
 Query Match 38.8%; Score 459; DB 13; Length 616;
 Best Local Similarity 99.8%; Pred. No. 2.7e-221;
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 9 CTCGAAGGACCCAGCCCTGCTCTGTGACAGAGCTCAAGGGCCCTGGCCCTCCCT 68
 DB 58 CTCGAAGGACCCAGCCCTGCTCTGTGACAGAGCTCAAGGGCCCTGGCCCTCCCT 117
 QY 69 CCTGGCTCGGCTGTGCTGGAGGGTTCCTCCAGTCCAGATCCCTAAGGAGCATGGGC 128
 DB 118 CCTGGCTCGGCTGTGCTGGAGGGTTCCTCCAGTCCAGATCCCTAAGGAGCATGGGC 177
 QY 129 AGCTGATCCATCCCTGGTGTACAACTGTGACTGACAGACAGATGCTAGCTACCCAAAC 188
 DB 178 AGCTGATCCATCCCTGGTGTACAACTGTGACTGACAGACAGATGCTAGCTACCCAAAC 237
 QY 189 CAACACTAGCCCTCCCTGAGATCCCTCCAGGCTGAGAGAGTCTGGGTGTCTCTAGGA 248
 DB 238 CAACACTAGCCCTCCCTGAGATCCCTCCAGGCTGAGAGAGTCTGGGTGTCTCTAGGA 297
 QY 249 CCAAGGACACTGGCAGACTTCAGAGAGGCCCCCAAGGCCCTAACTGTCCAGCCAGAGC 308
 DB 298 CCAAGGACACTGGCAGACTTCAGAGAGGCCCCCAAGGCCCTAACTGTCCAGCCAGAGC 357
 QY 309 ATCGGTCTCAGCAGAGTGTCTTCCCAAGCCTTGTATGACAAACCAATTTCCCTCGATGA 368
 DB 358 ATCGGTCTCAGCAGAGTGTCTTCCCAAGCCTTGTATGACAAACCAATTTCCCTCGATGA 417
 QY 369 TGTGCTCTGAGTGTCTGTGAGGAACAATGGAAAGTCTGCCAGCAGAGAAAATCTC 428

418 TGTGCTTCTGAGTCTGCTGTGAGGAAACAATGGGAAGTCTGCCAGCAGAGAAAATCTC 477
 QY 429 TGCAGAGCCCAAGCTTGTGAGTCTGCTGTCCAGGCGCAGGAGCTGTGACCATGGAGCAG 488
 DB 478 TGCAGAGCCCAAGCTTGTGAGTCTGCTGTCCAGGCGCAGGAGCTGTGACCATGGAGCAG 537
 QY 489 AGAGAGCAAGGCGCACAGCCGTGGCCCTGG 518
 DB 538 AGAGAGCAAGGCGCACAGCCGTGGCCCTGG 567
RESULT 7
LOCUS BG178487 778 bp mRNA linear EST 06-FEB-2001
DEFINITION 602328305F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429896 5',
 mRNA sequence.
ACCESSION BG178487
VERSION BG178487.1 GI:12685190
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 778)
COMMENT NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: DCTD/DTp
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNA10182 row: i column: 01
 High quality sequence stop: 657.
FEATURES
 Location/Qualifiers
 1..778
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4429896"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
ORIGIN
 Query Match 36.0%; Score 426; DB 12; Length 778;
 Best Local Similarity 99.7%; Pred. No. 1.7e-204;
 Matches 596; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 62 CTTCCCTCCCTGGCTCGGCTGTGCTTGGAGGGTCCCGAGTCCAGATCCCTAAGGAGC 121
 DB 26 CTTCCCTCCCTGGCTCGGCTGTGCTTGGAGGGTCCCGAGTCCAGATCCCTAAGGAGC 85
 QY 122 ATGGGGCAGCTGATCCATCCCTGCTGTACAACTGTGACTGCAGACAGATGCTGAGCTA 181
 DB 86 ATGGGGCAGCTGATCCATCCCTGCTGTACAACTGTGACTGCAGACAGATGCTGAGCTA 145
 QY 182 CCCAAACCAACCTAGCCTCTCCCTGAAGATCCTCCAGGCTCAGAGAGTTCTGGGTGT 241
 DB 146 CCCAAACCAACCTAGCCTCTCCCTGAAGATCCTCCAGGCTCAGAGAGTTCTGGGTGT 205
 QY 242 CCTAGGACCAAGGACACTGGCAGCTTCCAGAGGGCCCCCAAGCCCTACCTGTCCAG 301
 DB 206 CCTAGGACCAAGGACACTGGCAGCTTCCAGAGGGCCCCCAAGCCCTACCTGTCCAG 265

QY 302 CCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCC 361
 Db 266 CCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCC 325
 QY 362 TCATGATGCTCTCTGAGTGTCTCTGAGGACCAATGGGAGTCTGCCAGCAGAGA 421
 Db 326 TCATGATGCTCTCTGAGTGTCTCTGAGGACCAATGGGAGTCTGCCAGCAGAGA 385
 QY 422 AATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTCCTCAAGGCCAGGACCTGTGACCATG 481
 Db 386 AATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTCCTCAAGGCCA-GGACCTGTGACCATG 444
 QY 482 GAAGCAGAGAGAGCAAGCCGACAGCCGTGGCCCTGGGAGTCTTCCGGCAGGTGGCCG 541
 Db 445 GAAGCAGAGAGAGCAAGCCGACAGCCGTGGCCCTGGGAGTCTTCCGGCAGGTGGCCG 504
 QY 542 GCGAGCTGTGCTGAGACTCGGGGAGCCATTGACCATCTCTCTGAGGATGGAGACTGG 601
 Db 505 GCGAGCTGTGCTGAGACTCGGGGAGCCATTGACCATCTCTCTGAGGATGGAGACTGG 564
 QY 602 TGACCGTGTCTGTAATCTCAGCAGAGAGTATAACATCCCGAGGTCACAGTGG 659
 Db 565 TGACCGTGTCTGTAATCTCAGCAGAGAGTATAACATCCCGAGGTCACAGTGG 622

RESULT 8
 BG284179
 LOCUS 602408226F1 NIH_MGC_91 566 bp mRNA linear EST 21-FEB-2001
 DEFINITION mRNA sequence.
 ACCESSION BG284179.1 GI:13034866
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 566)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10418 row: c column: 07
 High quality sequence start: 2
 High quality sequence stop: 566
 Location/Qualifiers
 1..566
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4520382"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

FEATURES
 source

1..566
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4520382"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 34.2%; Score 405; DB 12; Length 566;
 Best Local Similarity 100.0%; Pred. No. 6.8e-194;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 779 GGCTCTTACTCTCTGTCTCAGTCCGCTCAGCGGCCCTGCATCTCTGGGACCGGATCAGACAC 838
 Db 13 GGCTCTTACTCTCTGTCTCAGTCCGCTCAGCGGCCCTGCATCTCTGGGACCGGATCAGACAC 72
 QY 839 TACAGGATTCACATGCTCTGACATGGCTGTGTACATCTCACCGGCTCTCACTTCCCC 898
 Db 73 TACAGGATTCACATGCTCTGACATGGCTGTGTACATCTCACCGGCTCTCACTTCCCC 132
 QY 899 TCATCTCAGGCCCCGTGGACCATTAATCTGAGTGGCGGATGACATCTGCTGCTTACTC 958
 Db 133 TCATCTCAGGCCCCGTGGACCATTAATCTGAGTGGCGGATGACATCTGCTGCTTACTC 192
 QY 959 AAGGAGCCCTGTGTCTGTCAGAGGGTGGCCCGCTCCCTGGCAAGGATATACCCCTACCT 1018
 Db 193 AAGGAGCCCTGTGTCTGTCAGAGGGTGGCCCGCTCCCTGGCAAGGATATACCCCTACCT 252
 QY 1019 GTGACTGTGAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTGTTTCT 1078
 Db 253 GTGACTGTGAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTGTTTCT 312
 QY 1079 GAAGTCTCCACAGGGAGGAGTCTCTTCTCAGTGGAGTCTCCGGAGTCCCTCAGCTTC 1138
 Db 313 GAAGTCTCCACAGGGAGGAGTCTCTTCTCAGTGGAGTCTCCGGAGTCCCTCAGCTTC 372
 QY 1139 TACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCCTAG 1183
 Db 373 TACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCCTAG 417

RESULT 9
 BQ053486
 LOCUS 6053486
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ053486.1 GI:19812826
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 878)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLCM2122 row: 1 column: 06
 High quality sequence stop: 394.
 Location/Qualifiers
 1..878
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5935253"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_106"
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a

FEATURES
 source

1..878
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5935253"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_106"
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a

db_xref="taxon:9606"
/map="20"
/clone_lib="pool_YT_lib_v_SPD"

ORIGIN

Query Match 25.4%; Score 300; DB 9; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.4e-140;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 CGCCTCACCTTCCCTCACTCCAGGCCCTCGTCCCTGAGAGGCTGGCCCGCTCCCTGSCAAG 1003
DB 37 CGCCTCACCTTCCCTCACTCCAGGCCCTCGTCCCTGAGAGGCTGGCCCGCTCCCTGSCAAG 96
QY 944 ATCTGCTGCTACTCAAGGAGCCCTGTCCTGTCAGAGGCTGGCCCGCTCCCTGSCAAG 1003
DB 97 ATCTGCTGCTACTCAAGGAGCCCTGTCCTGTCAGAGGCTGGCCCGCTCCCTGSCAAG 156
QY 1004 GATATACCCCTACCTCTGACTGTGACAGAGGACACCACTCACTGGAAGAGCTGGACAG 1063
DB 157 GATATACCCCTACCTCTGACTGTGACAGAGGACACCACTCACTGGAAGAGCTGGACAG 216
QY 1064 TCCTCTCTGTTTCTCAAGCTGCCAGGGAGGAGTCTCTTCTCAGTCAGGCTCTCCGG 1123
DB 217 TCCTCTCTGTTTCTCAAGCTGCCAGGGAGGAGTCTCTTCTCAGTCAGGCTCTCCGG 276
QY 1124 GAGTCCTCTCAGCTTCTACATCAGCTGAATGACAGGCTGTCTTTGATGATGCTAG 1183
DB 277 GAGTCCTCTCAGCTTCTACATCAGCTGAATGACAGGCTGTCTTTGATGATGCTAG 336

RESULT 11

AL844307 614 bp mRNA linear EST 30-JUL-2002
LOCUS
DEFINITION AL844307 pool_YT_lib_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION AL844307
VERSION AL844307.1 GI:22019089
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 614)
AUTHORS Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and Sheridan,E.
TITLE Homo sapiens EST sequence
JOURNAL Unpublished (2002)
COMMENT Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquerry@sanger.ac.uk
Sanger Centre name : sccdi0816.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool_YT_lib_v_SPD cDNA library. Further information can be found at
http://www.sanger.ac.uk/Teams/Team69/.

FEATURES
Location/Qualifiers
1..614
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="20"
/clone_lib="pool_YT_lib_v_SPD"

ORIGIN

Query Match 23.3%; Score 276; DB 9; Length 614;
Best Local Similarity 100.0%; Pred. No. 2.2e-128;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 908 GCCCTGGTGGACCATCTACTCTGAGCTGGCGGATGATCTGCTGCTACTCAAGAGGCC 967

NIH_MGC Library."

ORIGIN

Query Match 33.0%; Score 390; DB 12; Length 878;
Best Local Similarity 99.8%; Pred. No. 3.3e-186;
Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 179 CTACCCAAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGGTTCTGGG 238
DB 54 CTACCCAAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGGTTCTGGG 113
QY 239 TGTCTTAGACCAAGGACACTGGAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTC 298
DB 114 TGTCTTAGACCAAGGACACTGGAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTC 173
QY 299 CAGCAGAGCATGGCTCTCAGCAGAGCTGTCTTCCCAAGCCCTTTGATGACAAACCAATTT 358
DB 174 CAGCAGAGCATGGCTCTCAGCAGAGCTGTCTTCCCAAGCCCTTTGATGACAAACCAATTT 233
QY 359 CCCTCGATGATGTGCTTCTGAGTGTCTGCTGAGGACAAATGGGAAGTCTGCCAGCAGA 418
DB 234 CCCTCGATGATGTGCTTCTGAGTGTCTGCTGAGGACAAATGGGAAGTCTGCCAGCAGA 293
QY 419 AGAAATCTTGCCTGAGGAGGAGGAGTGTGCTTCTGCTGAGGAGGAGGAGTGTGACC 478
DB 294 AGAAATCTTGCCTGAGGAGGAGGAGTGTGCTTCTGCTGAGGAGGAGGAGTGTGACC 353
QY 479 ATGGAAGCAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 538
DB 354 ATGGAAGCAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 413
QY 539 CGGCGGAGAGTCTCCTGAGACTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 598
DB 414 CGGCGGAGAGTCTCCTGAGACTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 473

QY 599 TGTGAGGAGGAGTCTGCTGAA 619
DB 474 TGTGAGGAGGAGTCTGCTGAA 494

RESULT 10

AL844311 597 bp mRNA linear EST 30-JUL-2002
LOCUS
DEFINITION AL844311 pool_YT_lib_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION AL844311
VERSION AL844311.1 GI:22019093
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 597)
AUTHORS Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and Sheridan,E.
TITLE Homo sapiens EST sequence
JOURNAL Unpublished (2002)
COMMENT Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquerry@sanger.ac.uk
Sanger Centre name : sccdi0818.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool_YT_lib_v_SPD cDNA library. Further information can be found at
http://www.sanger.ac.uk/Teams/Team69/.

FEATURES
Location/Qualifiers
1..597
/organism="Homo sapiens"
/mol_type="mRNA"

```

Db      44  GCCCTGTGTGACCACTTCTGAGCTGGCGGATGACATCTGTCCTTACTCAAGGAGCCG 103
Qy      968  TGTGTCTCTGAGAGGGCTGCGCCGCTCCCTGGAAGGATATACCCCTACTGTGACTGTG 1027
Db      104  TGTGTCTCTGAGAGGGCTGCGCCGCTCCCTGGAAGGATATACCCCTACTGTGACTGTG 163
Qy      1028  CAGAGGACACACTCACTGGAAGAGCTGAGAGCTCCCTCTGTTTCTGAAGCTGCC 1087
Db      164  CAGAGGACACACTCACTGGAAGAGCTGAGAGCTCCCTCTGTTTCTGAAGCTGCC 223
Qy      1088  ACAGGGAGGAGTCTCTCTCACTGAGGCTCTCCGGAGTCTCCGGAGTCTCTACATCAGC 1147
Db      224  ACAGGGAGGAGTCTCTCTCACTGAGGCTCTCCGGAGTCTCCGGAGTCTCTACATCAGC 283
Qy      1148  CTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183
Db      284  CTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 319

```

```

RESULT 12
LOCUS   BG677567
DEFINITION BG677567 794 bp mRNA linear EST 01-MAY-2001
          602624118F1 NCI_CGAP_skn4 Homo sapiens cDNA clone IMAGE:474884 5',
          mRNA sequence.
ACCESSION BG677567
VERSION   BG677567.1 GI:13908964
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          NIH-MGC http://mgs.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgabbs@mail.nih.gov
          Tissue Procurement: James Cleaver, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM10601 row: d column: 05
          High quality sequence stop: 790.

```

```

FEATURES
          source
            1..794
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:474884"
              /tissue_type="squamous cell carcinoma"
              /lab_host="DH10B (T1 phage-resistant)"
              /clone_lib="NCI_CGAP_skn4"
              /note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;
              Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
              Average insert size 1.5kb. Library constructed by Life
              Technologies. Note: this is a NCI_CGAP Library."

```

```

ORIGIN
Query Match 21.6%; Score 255; DB 12; Length 794;
Best Local Similarity 100.0%; Pred. No. 1.1e-117;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      929  GAGCTGGGGATGACATCTGTGCTTCAAGGAGCCCTGTCTCCTGAGAGGGCTGGC 988
Db      17   GAGCTGGGGATGACATCTGTGCTTCAAGGAGCCCTGTCTCCTGAGAGGGCTGGC 76
Qy      989  CCGTCTCCCTGGAAGGATATACCCCTACTGTGACTGTGAGAGGACACCACTCACTGG 1048
Db      77   CCGTCTCCCTGGAAGGATATACCCCTACTGTGACTGTGAGAGGACACCACTCACTGG 136

```

```

Qy      1049  AAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCAGGGGAGGAGTCTCTTCTC 1108
Db      137  AAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCAGGGGAGGAGTCTCTTCTC 196
Qy      1109  AGTAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCTTCTACATCAGCTTCTCT 1168
Db      197  AGTAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCTTCTACATCAGCTTCTCT 256
Qy      1169  TTGGATGATGCTAG 1183
Db      257  TTGGATGATGCTAG 271

```

```

RESULT 13
LOCUS   BU944126
DEFINITION AGENCOURT 10545003 NIH_MGC_107 Homo sapiens cDNA clone
          IMAGE:6728350 5', mRNA sequence.
ACCESSION BU944126
VERSION   BU944126.1 GI:24132945
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          NIH-MGC http://mgs.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgabbs@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLC3049 row: m column: 21
          High quality sequence stop: 628.

```

```

FEATURES
          source
            1..960
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:6728350"
              /tissue_type="adenocarcinoma, cell line"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_107"
              /note="Organ: Breast; Vector: pOTB7; Site 1: EcoRI;
              Site 2: XhoI; cDNA made by oligo-dT priming.
              Directionally cloned into EcoRI/XhoI sites using the
              following 5' adaptor: GGCACGAG(G). Library constructed by
              Ling Hong in the laboratory of Gerald M. Rubin (University
              of California, Berkeley) using ZAP-cDNA synthesis kit
              (Stratagene) and Superscript II RT (Life Technologies).
              Note: this is a NIH_MGC Library."

```

```

ORIGIN
Query Match 17.3%; Score 205; DB 13; Length 960;
Best Local Similarity 100.0%; Pred. No. 2.8e-92;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      979  GAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACTGTGACTGTGAGGAGCACC 1038
Db      177  GAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACTGTGACTGTGAGGAGCACC 236
Qy      1039  ACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGGA 1098
Db      237  ACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGGA 296
Qy      1099  GTCTCTTCTCAGTGGGGTCTCCGGAGTCCCTCACTTCTTACATCCTTCAATCCTCA 1158

```

Db	297	GTCTCTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGA	356
Qy	1159	GCCTGTCTCTTTGGATGATGCCTAG	1183
Db	357	GGCTGTCTCTTTGGATGATGCCTAG	381
RESULT 14			
AL844309			
LOCUS	AL844309	pool_FLU_lib_v_SPC Homo sapiens	605 bp mRNA linear EST 30-JUL-2002
DEFINITION	AL844309		cdNA, mRNA sequence.
ACCESSION	AL844309.1	GI:22019091	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
		Homo sapiens (human)	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE			
AUTHORS		Ashecroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and Sheridan, E.	
TITLE		Homo sapiens EST sequence	
JOURNAL		Unpublished (2002)	
COMMENT		Contact: The Sanger Centre The Sanger Centre Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK Email: humquerry@sanger.ac.uk Sanger Centre name: sccdi0817.1541364 Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool FLU lib v_SPC cDNA library. Further information can be found at http://www.sanger.ac.uk/Teams/Team69/.	
FEATURES			
source		1..605 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /map="20" /clone_lib="pool_FLU_lib_v_SPC"	
ORIGIN			
Query Match		16.9%	Score 200; DB 9; Length 606;
Best Local Similarity		99.6%	Pred. No. 8.2e-90;
Matches		250; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	908	GCCTGTGTGACCACTTACTCTGAGCTGCGGATGACATCTCTGCCTACTCAAGAGCCC	967
Db	46	GCCTGTGTGACCACTTACTCTGAGCTGCGGATGACATCTCTGCCTACTCAAGAGCCC	105
Qy	968	TGTGTCTCTCAGAGGGTGGCCCGCTCCCTGGCAAGATATACCCCTACCTGTGACTGTG	1027
Db	106	TGTGTCTCTCAGAGGGTGGCCCGCTCCCTGGCAAGATATACCCCTACCTGTGACTGTG	165
Qy	1028	CAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCCCTGTTTCTGAAAGCTGCC	1087
Db	166	CAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCCCTGTTTCTGAAAGCTGCC	225
Qy	1088	ACAGGGGAGAGTCTTCTCTCAGTGAGGGTTCGGGAGTCCCTCAGCTTCTCATCAGC	1147
Db	226	ACAGGGGAGAGTCTTCTCTCAGTGAGGGTTCGGGAGTCCCTCAGCTTCTCATCAGC	285
Qy	1148	CTGAATGACGA	1158
Db	286	CTGAATGACGA	296
RESULT 15			
AL844309			

are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool_FLU_lib_v_SPC cDNA library. Further information can be found at <http://www.sanger.ac.uk/teams/team69/>.

FEATURES

source

1..611
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="20"
/clone_lib="pool_FLU_lib_v_SPC"

ORIGIN

Query Match 9.0%; Score 106; DB 9; Length 611;
Best Local Similarity 100.0%; Pred. No. 4.4e-42;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1078 TGAAGCTGCCACAGGGAGGAGTCTTCTCAGTGAGGCTTCCGGAGTCCCTCAGCTT 1137
DB 577 TGAAGCTGCCACAGGGAGGAGTCTTCTCAGTGAGGCTTCCGGAGTCCCTCAGCTT 518
QY 1138 CTACATCAGCTGATGACGAGGCTGCTCTTTGGATGATGCTAG 1183
DB 517 CTACATCAGCTGATGACGAGGCTGCTCTTTGGATGATGCTAG 472

RESULT 17

B45150

LOCUS

B45150 506 bp DNA linear GSS 21-OCT-1997
HS-1060-B1-H07-MF-abi CIT Human Genomic Sperm Library C Homo sapiens genomic clone Plate=CT 782 Col=13 Row=P, genomic survey sequence.

ACCESSION

B45150

VERSION

B45150.1

GI:2549984

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackrone@u.washington.edu
Sequence tagged Connector
Plate: CT 782 row: P column: 13
Class: BAC ends
High quality sequence stop: 506.
Location/Qualifiers
1..506
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Plate=CT 782 Col=13 Row=P"
/sex="M"
/clone_lib="CIT Human Genomic Sperm Library C"
/note="Origin: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

FEATURES

source

1..506
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Plate=CT 782 Col=13 Row=P"
/sex="M"
/clone_lib="CIT Human Genomic Sperm Library C"
/note="Origin: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 8.7%; Score 103; DB 28; Length 506;
Best Local Similarity 100.0%; Pred. No. 1.4e-40;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 AGAGAGAAGCAAGGCCACAGCCGTGGCCCTGGGAGTTTCCCGGAGGTGGCCGCCGCA 546
DB 210 AGAGAGAAGCAAGGCCACAGCCGTGGCCCTGGGAGTTTCCCGGAGGTGGCCGCCGCA 269
QY 547 GCTGCTGCTGAGACTCGGGAGGCAATGACCATCGTCTCTGAG 589
DB 270 GCTGCTGCTGAGACTCGGGAGGCAATGACCATCGTCTCTGAG 312

RESULT 18

AL844312/c

LOCUS

AL844312 pool_YT_lib_v_SPD Homo sapiens cDNA, mRNA linear EST 30-JUL-2002

DEFINITION

ACCESSION

AL844312

VERSION

AL844312.1

GI:22019094

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 619)

AUTHORS

TITLE

JOURNAL

COMMENT

Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquery@sanger.ac.uk
Sanger Centre name: sccdi0818.400489A
Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool_YT_lib_v_SPD cDNA library. Further information can be found at <http://www.sanger.ac.uk/teams/team69/>.

FEATURES

source

1..619
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="20"
/clone_lib="pool_YT_lib_v_SPD"

ORIGIN

Query Match 6.8%; Score 80; DB 9; Length 619;
Best Local Similarity 100.0%; Pred. No. 6.9e-29;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1068 TCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTTCTCAGTGAGGCTTCCGGAGT 1127
DB 570 TCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTTCTCAGTGAGGCTTCCGGAGT 511
QY 1128 CCCTCAGCTTCTACATCAGC 1147
DB 510 CCCTCAGCTTCTACATCAGC 491

RESULT 19

AQ556467

LOCUS

AQ556467 546 bp DNA linear GSS 29-MAY-1999

DEFINITION

ACCESSION

AQ556467

VERSION

AQ556467.1

GI:4916199

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 546)

AUTHORS

TITLE

JOURNAL

COMMENT

Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

Email: humquery@sanger.ac.uk

Sanger Centre name: sccdi0818.400489A

Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool_YT_lib_v_SPD cDNA library. Further information can be found at http://www.sanger.ac.uk/teams/team69/.

AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Kellner,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

PUBMED 10449764

COMMENT Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu Plate: 812 row: F column: 1 Seq primer: T7 Class: BAC ends High quality sequence stop: 546. Location/Qualifiers 1..546 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone="pPlate=812 Col=1 Row=F" /sex="male" /clone_lib="RPCI-11 Human Male BAC Library" /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

FEATURES source Query Match 5.7%; Score 68; DB 28; Length 546; Best Local Similarity 100.0%; Pred. No. 8.2e-23; Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN 355 ATTTCCTCGATGATGCTTCTGATGCTCTGCTGAGGACAAATGGGAAGTCTGCCAG 414 Db 383 ATTTCCTCGATGATGCTTCTGATGCTCTGCTGAGGACAAATGGGAAGTCTGCCAG 442 QY 415 CAGAAGAA 422 Db 443 CAGAAGAA 450

RESULT 20 AA309769 149 bp mRNA linear EST 19-APR-1997 LOCUS BSL180699 Jurkat T-cells V Homo sapiens cDNA 5' end, mRNA sequence. DEFINITION AA309769 ACCESSION AA309769.1 GI:1962171 VERSION EST. KEYWORDS Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 149) Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodak,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelsey,J.C., Liu,I.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,

Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,J., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseitine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE 96026280

PUBMED 7566098

COMMENT Contact: Kerlavage, AR Bioinformatics The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850 USA Tel: 3018699056 Fax: 3018699423 Email: arkerlav@tigr.org For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tdb/hgi/hgi.html) Seq primer: M13 Reverse. Location/Qualifiers 1..149 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="ATCC (inhost):155929" /db_xref="taxon:9606" /clone_lib="T-lymphocyte" /clone="T-Jurkat T-cells V" /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

FEATURES source Query Match 4.8%; Score 57; DB 9; Length 149; Best Local Similarity 100.0%; Pred. No. 1.9e-17; Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN 1127 TCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTCTTGGATGATGCTAG 1183 Db 1 TCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTCTTGGATGATGCTAG 57

RESULT 21 AQ556478 553 bp DNA linear GSS 29-MAY-1999 LOCUS HS_5236 B1.D02.T7A RPCI-11 Human Male BAC Library Homo sapiens genomic_clone Plate=812 Col=3 Row=H, genomic survey sequence. DEFINITION AQ556478 ACCESSION AQ556478.1 GI:4916210 VERSION GSS. KEYWORDS Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 553) Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Kellner,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

PUBMED 10449764

COMMENT Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887

		/tissue_type="retina"
		/dev_stage="adult"
		/lab_host="DH10B"
		/clone_lib="RIKEN full-length enriched, adult retina"
		/notes=Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGGATCCCAAGGTCTTTTTTTTWTN 3'], cDNA was
		GAGAGAGAGGATCCCAAGGTCTTTTTTTTWTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second
		strand cDNA was prepared with the primer adapter of sequence [5',GAGAGAGAGATTCTCGAGTAATAAATTAATACCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. -Retina RNA was provided by Stefano Gustinich, Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA02115, USA, whose assistance we gratefully acknowledge."
ORIGIN		
Query Match	3.6%;	Score 43; DB 13; Length 660;
Best Local Similarity	100.0%;	Pred. No. 4.3e-10;
Matches	43; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
DY	740	GGAGGGGCCTTCCTCATCCGGGAGGCCAGACCAGGAGGCT 782
Ddb	227	GGAGGGGCCTTCCTCATCCGGGAGGCCAGACCAGGAGGCT 269
<hr/>		
RESULT 24		
AKO20837		
LOCUS		
DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930009SE21 product:CDNATOR OF ANTIGEN RECEPTOR SIGNALING MAPS, full insert sequence.	AKO20837 326 bp mRNA linear HTC 20-SEP-2003
ACCESSION	AKO20837.1	GI:12861542
VERSION	HTC; CAP trapper.	
KEYWORDS	Mus musculus (house mouse)	
SOURCE	Mus musculus	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapped-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
AUTHORS	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapped-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
TITLE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		

JOURNAL REFERENCE AUTHORS	Nature 420, 563-573 (2002) 6 (bases 1 to 2637)
LOCUS	AK030877
DEFINITION	Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:583043/K10 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARKS, full insert sequence.
ACCESSION	AK030877.1 GI:26326848
VERSION	AK030877
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus (house mouse)
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
TITLE	Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishiguchi, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Oikawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
TITLE	Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	11076861
PUBMED	11076861
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishiguchi, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Oikawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
TITLE	Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	11076861
PUBMED	11076861
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishiguchi, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Oikawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
TITLE	Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	11076861
PUBMED	11076861
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishiguchi, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Oikawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
TITLE	Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	11076861
PUBMED	11076861
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishiguchi, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki,

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

1. .2974
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:5830437K10"
/db_xref="MGI:2394046"
/db_xref="taxon:10090"
/clone="5830437K10"
/sex="male"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
232..1011

CDS

/note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTR|AAL38196, evidence: FASTY, 100%ID, 100%length, match=777)
putative"
/codon_start=1
/protein_id="BAC27168.1"
/db_xref="GI:26326849"
GQARLSURGLPLTII SEDGWTQVSEGVREYHMPVYVAKVHGLVGLSREK
ABELLLPFGPGALRSGYSLVRLSPASWDRIRHRIORLONGLWLYIT
PLRTPSLHALVHVSLEADGICCPRLRCPVLQKGLPGLKDTPEPTVPVTSLSNWK
LDRSLFLFAPASGASLILSEGLRESLSYISLAEDPLDDA"

ORIGIN

Query Match 3.6%; Score 43; DB 11; Length 2974;
Best Local Similarity 100.0%; Pred. No. 7.6e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTCATCCGGAGAGCCAGACGAGAGGCT 782

Db 571 GGAGGGGCTTCTCATCCGGAGAGCCAGACGAGAGGCT 613

RESULT 27

AW437301

LOCUS

AW437301 322 bp mRNA linear EST 25-APR-2001

DEFINITION 78266 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION AW437301

VERSION AW437301.1 GI:6972607

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

1 (bases 1 to 322)

Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Casas, E., Wray, J.E., White, J., Cho, J., Fahrner, S.C.,

Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,

Chitko-McKown, C.G., Perte, G., Holt, I., Karamycheva, S., Liang, F.,

Quackenbush, J. and Keele, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

11282978

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCAGTCACGACG

Plate: 40 row: K column: 1

Seq primer: ATTATGCTGACACTATAG.

FEATURES

source

1. .322
Location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

ORIGIN

Query Match 3.6%; Score 42; DB 10; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1033 GACACCACTCACTGGAAGAGCTGGACAGCTCCCTCCTGTT 1074

Db 4 GACACCACTCACTGGAAGAGCTGGACAGCTCCCTCCTGTT 45

RESULT 28

CB426333/c

LOCUS

CB426333 569 bp mRNA linear EST 25-MAR-2003

DEFINITION 601508 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.

ACCESSION CB426333

VERSION CB426333.1 GI:29198274

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

1 (bases 1 to 569)

Smith, T.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-McKown, C.G.,

Wray, J.E. and Keele, J.W.

A second set of bovine ESTs from pooled-tissue normalized libraries

Unpublished (2003)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and

trimmed with the aid of the trim_alt option. Vector identified with

cross_match v0.990329.

Plate: PQY8029 row: E column: 5

Seq primer: TAGAAGGCACAGTCGAGG.

FEATURES

source

1. .569
Location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

ORIGIN

Query Match 3.6%; Score 42; DB 14; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

MEDLINE
PUBMED
COMMENT

22354164
12466305
Other ESTs: H4019E02-5
Contact: Yong Qian
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
Email: cdna@gsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://gsun.grc.nia.nih.gov/cdna/NIA_7_4k.html for details.
Plate: H4019 row: E column: 02
Seq primer: -21M13 Forward
High quality sequence stop: 603
POLYA=Yes.

FEATURES

source

Location/Qualifiers

1. .603
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="niaEST:H4019E02-3"
/db_xref="taxon:10090"
/clone="H4019E02"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."

ORIGIN

Query Match 2.5%; Score 29; DB 13; Length 603;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CTTGACATGGCTGGCTGACATCTCACC 882

DB 588 CTTGACATGGCTGGCTGACATCTCACC 560

RESULT 32

BB619854

LOCUS BB619854 RIKEN full-length enriched, adult male thymus Mus musculus
DEFINITION cDNA clone 5830437K10 5', mRNA sequence.
ACCESSION BB619854
VERSION BB619854.1 GI:15396929

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 627)

REFERENCE

AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, H., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)

TITLE

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saitama-shi, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
Wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A.,
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K.,
Arakawa, T., Ishii, Y. and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
Funct. Genomics 2 pre, L72-L86 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

e mouse tissues.

FEATURES

source

Location/Qualifiers

1. .627
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5830437K10"
/sex="male"
/tissue_type="thymus"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male thymus"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Ret = 10.0 and subtraction to Ret = 100.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTTAATAATTAATCCCCCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:
BamHI."

ORIGIN

Query Match 2.4%; Score 28; DB 10; Length 627;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 GGAGGGGCTTCTCTATCCGGAGAGACC 767

DB 568 GGAGGGGCTTCTCTATCCGGAGAGACC 595

RESULT 33

CC527971/c

LOCUS

DEFINITION

CC527971

ACCESSION

VERSION

KEYWORDS

SOURCE

CC527971 757 bp DNA linear GSS 17-JUN-2003
CH240_403E18.T7 CHORI-240 Bos taurus genomic clone CH240_403E18,
genomic survey sequence.
CC527971
CC527971.1 GI:31846259
GSS.
Bos taurus (cow)

```

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 757)
AUTHORS Holt, R., Stott, J., Yang, G., Barber, S., Smallus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Garn, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., Keele, J.W. and Kappes, S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_403E18.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 403 row: E column: 18
Seq primer: T7
Class: BAC ends.
FEATURES             source
    source
    1..757
        /organism="Bos taurus"
        /mol_type="genomic DNA"
        /strain="breed: Hereford"
        /db_xref="taxon:9913"
        /clone="CH240_403E18"
        /sex="Male"
        /cell_type="Blood"
        /clone_lib="CHORI-240"
        /note="Vector: pPARBAC1.3; Site 1: MboI; Site 2: MboI;
        Hereford bull LI Domino 99375; CHORI-240 Bovine BAC
        library (Male) produced by Pieter de Jong"
ORIGIN
    Query Match      2.3%; Score 27; DB 29; Length 757;
    Best Local Similarity 100.0%; Pred. No. 0.06;
    Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 589 GGATGGAGACTGGTGGACGGTGCTCTC 615
Db 297 GGATGGAGACTGGTGGACGGTGCTCTC 271

RESULT 34
B1898999
LOCUS 480839 MARC 2BOV Bos taurus cdna 5', mRNA linear EST 16-OCT-2001
DEFINITION B1898999
ACCESSION B1898999
VERSION B1898999.1 GI:16187055
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM
REFERENCE
AUTHORS Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,
Bennett, G.B., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,
Chitko-Mckown, C.G., Perle, G., Holt, I., Karameycheva, S., Liang, F.,
Quackenbush, J. and Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cdna
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
MEDLINE 11282978
PUBMED 11282978
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.990904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACAGTATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 105 row: P column: 20
Seq primer: ATTAGTGACACTATAG.
FEATURES             source
    source
    1..541
        /organism="Bos taurus"
        /mol_type="mRNA"
        /db_xref="taxon:9913"
        /tissue_type="pooled"
        /lab_host="DH10B"
        /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
        Library made from pooled tissue from testis, thymus,
        semitendinosus muscle, longissimus muscle, pancreas,
        adrenal, and endometrium."
ORIGIN
    Query Match      1.9%; Score 23; DB 12; Length 541;
    Best Local Similarity 100.0%; Pred. No. 5.7;
    Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 506 GCCGTGGCCCTGGCAGTTTCCC 528
Db 441 GCCGTGGCCCTGGCAGTTTCCC 463

RESULT 35
BG388843/c
LOCUS 602414563F1 NIH_MGC_92 Homo sapiens cdna clone IMAGE:4522958 5',
DEFINITION BG388843
ACCESSION BG388843
VERSION BG388843.1 GI:13282289
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS NTH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@email.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM10424 row: n column: 15
High quality sequence stop: 725.
FEATURES             source
    source
    1..764
        /organism="Homo sapiens"
        /mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:4522958"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="PH108 (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 1.9%; Score 22; DB 12; Length 764;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 TCAGCAGAGCTGTCTTCCCAAG 337
Db 492 TCAGCAGAGCTGTCTTCCCAAG 471

RESULT 36
LOCUS CB994934/c
DEFINITION AGENCOURT_1362391.8 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30335210 5', mRNA sequence.
ACCESSION CB994934
VERSION CB994934.1 GI:30289454
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 791)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM356 row: 1 column: 03
High quality sequence stop: 558.
Location/Qualifiers
1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30335210"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/notes="Organ: Placenta; Vector: pBluescriptB; Site 1:
all-XhoI; Site 2: BamHI; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 1.9%; Score 22; DB 14; Length 791;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 316 TCAGCAGAGCTGTCTTCCCAAG 337
Db 527 TCAGCAGAGCTGTCTTCCCAAG 506

RESULT 37
LOCUS BG282272/c
DEFINITION 602402951.F1 NIH_MGC_20 Homo sapiens cDNA clone
IMAGE:4545036 5',
mRNA sequence.
ACCESSION BG282272
VERSION BG282272.1 GI:113031199
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 875)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW1229 row: f column: 13
High quality sequence stop: 633.
Location/Qualifiers
1..875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4545036"
/tissue_type="melanotic melanoma"
/clone_lib="NIH_MGC_20"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Size-selected 500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
Query Match 1.9%; Score 22; DB 12; Length 875;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 TCAGCAGAGCTGTCTTCCCAAG 337
Db 512 TCAGCAGAGCTGTCTTCCCAAG 491

RESULT 38
LOCUS BI549288/c
DEFINITION 603189926.F1 NIH_MGC_95 Homo sapiens cDNA clone
IMAGE:5261143 5',
mRNA sequence.
ACCESSION BI549288
VERSION BI549288.1 GI:15436600
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 934)

```

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NMGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM11658 row: d column: 08
 High quality sequence stop: 710.
 Location/Qualifiers
 1. 934
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5261143"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /clone_lib="NTH MGC 95"
 /note="Organ: brain; Vector: pBluescriptR (modified Bluescript KS-); Site: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NMGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 1.9%; Score 22; DB 12; Length 934;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 316 TCAGCAGAGCTGCTTCCCAAG 337
 |||||
 Db 512 TCAGCAGAGCTGCTTCCCAAG 451
 |||||

RESULT 39
 BM926459/c
 LOCUS 1109 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT_6644584 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5766843 5', mRNA sequence.
 ACCESSION BM926459
 VERSION BM926459.1 GI:19376838
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1109)
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM12824 row: c column: 04
 High quality sequence stop: 696.
 Location/Qualifiers

FEATURES
 source
 Query Match 1.9%; Score 22; DB 12; Length 934;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1. 1109
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5766843"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_122"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dr primed, and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 1.9%; Score 22; DB 12; Length 1109;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 316 TCAGCAGAGCTGCTTCCCAAG 337
 |||||
 Db 616 TCAGCAGAGCTGCTTCCCAAG 595
 |||||

RESULT 40
 AY418068/c
 LOCUS 2517 bp DNA linear GSS 17-DEC-2003
 DEFINITION Homo sapiens OCA2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
 ACCESSION AY418068
 VERSION AY418068.1 GI:39774028
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2517)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 2517)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering them based on alignment.
 Location/Qualifiers
 1. 2517
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..2517
 /gene="OCA2"
 /locus_tag="HCM6434"

ORIGIN
 Query Match 1.9%; Score 22; DB 29; Length 2517;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 316 TCAGCAGAGCTGCTTCCCAAG 337
Db 398 TCAGCAGAGCTGCTTCCCAAG 377

RESULT 41
AY418069/c 2517 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes OCA2 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY418069
VERSION AY418069.1 GI:39774029
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
REFERENCE
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location/Qualifiers
1..2517
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2517
/gene="OCA2"
/locus_tag="HCM6434"

ORIGIN
Query Match 1.9%; Score 22; DB 29; Length 2517;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 TCAGCAGAGCTGCTTCCCAAG 337
Db 398 TCAGCAGAGCTGCTTCCCAAG 377

RESULT 42
CB657651
LOCUS OSJNEC13C17.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEC13C17 5', mRNA sequence.
ACCESSION CB657651
VERSION CB657651.1 GI:29661376
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM
REFERENCE
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

QY 316 TCAGCAGAGCTGCTTCCCAAG 337
Db 398 TCAGCAGAGCTGCTTCCCAAG 377

RESULT 43
BE030537/c 332 bp mRNA linear EST 09-JUL-2000
LOCUS BE030537
DEFINITION BE030537
ACCESSION BE030537
VERSION BE030537.1 GI:8325546
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 332)
Fahrenkrug,S.C., Smith,T.P.L., Fekking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Perrea,G., Sultana,R.,
Quackenbush,J. and Keefe,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
JOURNAL 22213789
MEDLINE 12226715
PUBMED
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACACCTATGACCAT
BACKWARD: GTTTCACGTCACGAG
Plate: 62 row: 1 column: 14
Seq primer: ATTAGGTGACATATAG.
Location/Qualifiers

```

```

JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 13 row: C column: 17
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1..320
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC13C17"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEC"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with RiceBlast (C9240-1)"

ORIGIN
Query Match 1.8%; Score 21; DB 14; Length 320;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 GCCCGGCCGAGCTGCTGCTGA 557
Db 45 GCCCGGCCGAGCTGCTGCTGA 65

RESULT 43
BE030537/c 332 bp mRNA linear EST 09-JUL-2000
LOCUS BE030537
DEFINITION BE030537
ACCESSION BE030537
VERSION BE030537.1 GI:8325546
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 332)
Fahrenkrug,S.C., Smith,T.P.L., Fekking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Perrea,G., Sultana,R.,
Quackenbush,J. and Keefe,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
JOURNAL 22213789
MEDLINE 12226715
PUBMED
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACACCTATGACCAT
BACKWARD: GTTTCACGTCACGAG
Plate: 62 row: 1 column: 14
Seq primer: ATTAGGTGACATATAG.
Location/Qualifiers

```



```

source
1. .332
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

ORIGIN
Query Match 1.8%; Score 21; DB 10; Length 332;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 CTGACGAGGAGAGAGAGAGAG 712
|||||
Db 34 CTGACGAGGAGAGAGAGAGAG 14

RESULT 44
AA365070
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
COMMENT
AA365070 Pineal gland II Homo sapiens cDNA 5' end, mRNA sequence.
AA365070
AA365070.1 GI:2017387
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 363)
Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstein, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wei, C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S.,
Glocke, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Saudke, D.M., Shinkley, R.,
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A.,
He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,
Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
Fraser, C.M., and Venter, J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
7566098
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. .363
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):169794"
/db_xref="taxon:9606"
/dev_stage="adult"

FEATURES
source
1. .408
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129 OLA"
/db_xref="taxon:10090"
/sex="Male"
/cell_type="Embryonic Stem Cell"
/clone_lib="Sanger Institute Gene Trap Library PGT01xf"
/notes="Vector: pGT01xf"

ORIGIN
Query Match 1.8%; Score 21; DB 29; Length 408;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 477 CCATGGAAGCAGAGAGAGCA 497
|||||
Db 394 CCATGGAAGCAGAGAGAGCA 374

RESULT 46
CF540846/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
CF540846 454 bp mRNA linear EST 12-SEP-2003
UI-M-GWO-cio-1-21-0-UI.r1 NIH_BMAP_GWO Mus musculus cDNA clone
IMAGE:30542396 5', mRNA sequence.
CF540846
CF540846.1 GI:34593369
EST.
Mus musculus (house mouse)
Mus musculus
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 454)

FEATURES
source
1. .408
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129 OLA"
/db_xref="taxon:10090"
/sex="Male"
/cell_type="Embryonic Stem Cell"
/clone_lib="Sanger Institute Gene Trap Library PGT01xf"
/notes="Vector: pGT01xf"

ORIGIN
Query Match 1.8%; Score 21; DB 29; Length 408;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 477 CCATGGAAGCAGAGAGAGCA 497
|||||
Db 394 CCATGGAAGCAGAGAGAGCA 374

RESULT 46
CF540846/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
CF540846 454 bp mRNA linear EST 12-SEP-2003
UI-M-GWO-cio-1-21-0-UI.r1 NIH_BMAP_GWO Mus musculus cDNA clone
IMAGE:30542396 5', mRNA sequence.
CF540846
CF540846.1 GI:34593369
EST.
Mus musculus (house mouse)
Mus musculus
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 454)

```

AUTHORS **TITLE** **JOURNAL** **COMMENT**

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
(This clone was contributed by the Brain Molecular Anatomy Project (BMAP))

Seq primer: pYX-5.

FEATURES **source**

Location/Qualifiers
1. .454
/organism="Mus musculus"
/mol_type="rRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30542396"
/issue_type="whole eye"
/dev_stage="embryo 15.5, 16.5, 17.5, 18.5 dpc"
/lab_host="PH108 (T1 phase resistant)"
/clone_lib="NIH_BMAP_GW0"
/note="Organ: Eye; Vector: pYX-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CTCGCGCCCTC. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 1.8%; Score 21; DB 14; Length 454;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 477 CCATGAGCAGCAGAGAGCA 497
Db 110 CCATGAGCAGCAGAGAGCA 90

RESULT 47 **LOCUS** **DEFINITION**

AQ372680 561 bp DNA linear GSS 20-MAY-1999
RPC111-158D9.TV RPCI-11 Homo sapiens genomic clone RPCI-11-158D9,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
Unpublished (1997)

Other GSSs: RPC111-158D9.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES **source**

Location/Qualifiers
1. .561
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7560368"
/db_xref="taxon:9606"
/clone="RPCI-11-158D9"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; RPC111 Human Male BAC Library"

ORIGIN

Query Match 1.8%; Score 21; DB 28; Length 561;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1011 CCTACTGCTGACTGTGCAGA 1031
Db 213 CCTACTGCTGACTGTGCAGA 233

RESULT 48 **LOCUS** **DEFINITION**

AZ716247 653 bp DNA linear GSS 24-JAN-2001
RPCI-24-152E7.TV RPCI-24 Mus musculus genomic clone RPCI-24-152E7,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM
Mus musculus (house mouse)

REFERENCE
AUTHORS
Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M., Tsengave, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E., Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)

Other GSSs: RPCI-24-152E7.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://www.bacpac.org/bacpac/orderingframe.html>). BAC end search page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Place: 152 row: E column: 7
Seq primer: T7
Class: BAC ends.

FEATURES **source**

Location/Qualifiers
1. .653
/organism="Mus musculus"
/mol_type="genomic DNA"

```

/strain="C57BL/6J"
/db xref="taxon:10090"
/clone="RPC1-24-152E7"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPC1-24"
/notes=vector: PTABAC1, Site 1: BamHI, Site 2: BamHI;
RPC1-24 Mouse BAC library produced by Pieter de Jong. The
library was cloned in the PTABAC1 cloning vector C57BL/6J
BamHI sites using MboI partially digested male C57BL/6J
DNA."

ORIGIN
Query Match 1.8%; Score 21; DB 28; Length 653;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

477 CCATGGAAGCAGAGAGCA 497
|||||
337 CCATGGAAGCAGAGAGCA 317

RESULT 49
LOCUS BB629807.1 GI:15399698
DEFINITION Mus musculus (house mouse)
ACCESSION BB629807
VERSION BB629807.1
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Arakawa, T., Carrincci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ono, H., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
TITLE RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
JOURNAL
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carrincci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A.,
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K.,
Arakawa, T., Ishii, Y. and Hayashizaki, Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
Funct. Genomics 2 pre, L72-L86 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
FEATURES
Location/Qualifiers
1..668
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9930012K12"
/sex="female"
/tissue_type="vagina"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult female
vagina"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GACAGAGAAGCATCCAGAGCTCTTTTITTTTTTITVTN 3'], cDNA was

```

prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAATCCGCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 1.8%; Score 21; DB 10; Length 668;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 TCCAGAGAGGCCGCCCAAGCC 288
Db 402 TCCAGAGAGGCCGCCCAAGCC 382

RESULT 51

BI752759 821 bp mRNA linear EST 25-SEP-2001
LOCUS 603028385F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5198709 5',
DEFINITION mRNA sequence.

ACCESSION BI752759

VERSION BI752759.1 GI:15744337

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11497 row: j column: 22
High quality sequence stop: 732.

FEATURES

Source

1..821
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5198709"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yr. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.8%; Score 21; DB 12; Length 821;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

Db

106

274

AGATCCCTAAGGACATGGG 126

AGATCCCTAAGGACATGGG 294

RESULT 52

B202783/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 840)

Zhao,S., Shetty,J., Shatsman,S., Teegave,G., Geer,K.,

Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,

Riggs,F. de Jong,P. and Fraser,C.M.

Rat BAC End Sequences from Library CHORI-230 MboI segment

Unpublished (1999)

Other GSSs: CH230-351B11.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pdjong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or ering information.htm). BAC end

plate: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 351 row: B column: 11

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..840

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SSNHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-351B11"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 2"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by

Pieter de Jong"

ORIGIN

Query Match

Best Local Similarity

Matches

21; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

Db

1135

CTTCTACATCAGCTGAATGA 1155

CTTCTACATCAGCTGAATGA 360

CTTCTACATCAGCTGAATGA 360

CTTCTACATCAGCTGAATGA 360

CTTCTACATCAGCTGAATGA 360

CTTCTACATCAGCTGAATGA 360

CTTCTACATCAGCTGAATGA 360

BZ202783 840 bp DNA linear GSS 11-OCT-2002
CH230-351B11-TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-351B11, genomic survey sequence.

BZ202783

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 840)

Zhao,S., Shetty,J., Shatsman,S., Teegave,G., Geer,K.,

Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,

Riggs,F. de Jong,P. and Fraser,C.M.

Rat BAC End Sequences from Library CHORI-230 MboI segment

Unpublished (1999)

Other GSSs: CH230-351B11.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pdjong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or ering information.htm). BAC end

plate: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 351 row: B column: 11

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..840

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SSNHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-351B11"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 2"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by

Pieter de Jong"

ORIGIN

Query Match

Best Local Similarity

Matches

21; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

Db

2296 bp mRNA linear HTC 20-SEP-2003

Mus musculus adult male tongue cDNA, RIKEN full-length enriched

library, clone:2310001L23 product:similar to KERATIN, TYPE II

CYTOSKELETAL 2 ORAL (CYTOKERATIN 2P) (K2P) (CK 2P) [Homo sapiens],

full insert sequence.

AK009075

AK009075.1 GI:12843643

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
JOURNAL MEDLINE PUBMED

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
JOURNAL MEDLINE PUBMED

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
JOURNAL MEDLINE PUBMED

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861
JOURNAL MEDLINE PUBMED

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2296)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hirakawa, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Koike, Y., Komno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saitsu, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tegawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshiro, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAAGACCAAGCTCTTTTTCCTTTTTTTTNN 3'], cDNA was
prepared by using trihalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand was synthesized with the primer adapter of sequence [5'

```

REFERENCE
AUTHORS
  4 The RIKEN Genome Exploration Research Group Phase II Team and the
    FANTOM Consortium.
TITLE
  Functional annotation of a full-length mouse cDNA collection
REFERENCE
AUTHORS
  5 Nature 409, 685-690 (2001)
TITLE
  The FANTOM Consortium and the RIKEN Genome Exploration Research
  Group Phase I & II Team.
JOURNAL
  Analysis of the mouse transcriptome based on functional annotation
  of 60,770 full-length cDNAs
REFERENCE
AUTHORS
  6 (bases 1 to 2311)
  Nature 420, 563-573 (2002)
  Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
  Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
  Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
  Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
  Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
  Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
  Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
  Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
  Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
  Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
  Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
  Muramatsu, M. and Hayashizaki, Y.
  Direct Submission
  Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
  Physical and Chemical Research (RIKEN), Laboratory for Genome
  Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
  RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
  Kanagawa 230-0045, Japan (E-mail: genome-res@gscl.riken.go.jp,
  URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
  Fax: 81-45-503-9216)
COMMENT
  cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN.
  Division of Experimental Animal Research in Riken contributed to
  prepare mouse tissues.
  Please visit our web site for further details.
  URL: http://genome.gsc.riken.go.jp/
  URL: http://fantom.gsc.riken.go.jp/
FEATURES
  source
    1..2311
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6J"
      /db_xref="FANTOM DB:9930012K12"
      /db_xref="MGI:2401255"
      /db_xref="taxon:10090"
      /clone="9930012K12"
      /sex="female"
      /tissue_type="vagina"
      /clone_lib="RIKEN full-length enriched mouse cDNA library"
      /dev_stage="adult"
      63..1847
        /note="putative"
        similar to KERATIN, TYPE II CYTOSKELETAL 2 ORAL
        (CYTOKERATIN 2P) (K2P) (CK 2P) [Homo sapiens]
        (SWISSPROT Q01546, evidence: FASTY, 76%ID, 99.6%length,
        match=1779)"
        2291..2296
          /note="putative"
        2311
          /note="putative"
    misc_feature
      /note="putative"
      similar to KERATIN, TYPE II CYTOSKELETAL 2 ORAL
      (CYTOKERATIN 2P) (K2P) (CK 2P) [Homo sapiens]
      (SWISSPROT Q01546, evidence: FASTY, 76%ID, 99.6%length,
      match=1779)"
      2291..2296
        /note="putative"
      2311
        /note="putative"
    polyA_signal
      2291..2296
        /note="putative"
    polyA_site
      2311
        /note="putative"
  ORIGIN
    Query Match 1.8%; Score 21; DB 11; Length 2311;
    Best Local Similarity 100.0%; Pred. No. 1e+02;
    Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 268 TCCAGAGGGCCCCCAAGGCC 288
    DB 407 TCCAGAGGGCCCCCAAGGCC 387

```

```

RESULT 55
H75025/c
LOCUS       H75025       163 bp      mRNA      linear      EST 31-OCT-1995
DEFINITION 77 PtIFG1 Pinus taeda cDNA clone 1869s, mRNA sequence.
ACCESSION  H75025
VERSION    H75025.1   GI:1048500
KEYWORDS   EST.
SOURCE     Pinus taeda (loblolly pine)
ORGANISM   Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.
            1 (bases 1 to 163)
            Kinlaw, C.S.
            Loblolly pine cDNAs
            Unpublished (1995)
            Contact: Claire S. Kinlaw
            USDA IFG Dendrome Project
            Institute of Forest Genetics
            Dendrome Project, Institute of Forest Genetics, P.O. Box 245,
            Berkely, CA 94701
            Tel: 5105596429
            Fax: 5105596440
            Email: csk@sf27w007.pswfs.gov
            Seq primer: custom.
            Location/Qualifiers
              1..163
                /organism="Pinus taeda"
                /mol_type="mRNA"
                /db_xref="taxon:3352"
                /clone="1869s"
                /clone_lib="PtIFG1"
                /note="The tissue source for this library is seedlings.
                The cDNA was randomly primed and not directionally
                cloned."
  ORIGIN
    Query Match 1.7%; Score 20; DB 14; Length 163;
    Best Local Similarity 100.0%; Pred. No. 1.2e+02;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 412 CAGCAGAGAAATCTCTGC 431
    DB 49 CAGCAGAGAAATCTCTGC 30
RESULT 56
H75025/c
LOCUS       H75025       282 bp      mRNA      linear      EST 29-NOV-2000
DEFINITION BB563180 RIKEN full-length enriched, 18 days embryo Mus musculus
            cDNA clone 1100001B05 5', mRNA sequence.
ACCESSION  BB563180
VERSION    BB563180.1   GI:11454072
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
            1 (bases 1 to 282)
            Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T.,
            Carninci, P., Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T.,
            Hodojima, Y., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J.,
            Kojima, Y., Konno, H., Kusakabe, M., Matsuyama, T., Miyazaki, A.,
            Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Okazaki, Y.,
            Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Sato, K.,
            Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
            Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, T.,
            Watanabe, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A.,
            Muramatsu, M. and Hayashizaki, Y.
            RIKEN Mouse ESTs (Aizawa, K. et al. 2000)
            Unpublished (2000)
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic

```

Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermostable enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2) 520-524 (1998)
Itoh,M., Kiteunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

FEATURES
source
Location/Qualifiers
1..282
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="110001B05"
/dev_stage="18 days embryo"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, 18 days embryo"
/note="Site 1: XhoI; Site 2: SstI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGACAGAGCGCGCAACTCGAGTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGACAGAGAGGATCCAGAGCTCAATTAATTAAACCCCCCCCCC 3'].
cDNA was cleaved with XhoI and SstI."

Query Match 1.7%; Score 20; DB 10; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1076 TCTGAAGCTGCCACAGGGGA 1095
|||||
DB 98 TCTGAAGCTGCCACAGGGGA 117
|||||

RESULT 57
AA589496 341 bp mRNA linear EST 16-SEP-1997
LOCUS V148b12.61 Stratagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:975455 3' similar to gb:M61906 PHOSPHATIDYLINOSITOL 3-KINASE
REGULATORY ALPHA SUBUNIT (HUMAN); gb:M60651 Mouse 85 kDa
phosphoprotein (MOUSE); mRNA sequence.
ACCESSION AA589496
VERSION AA589496.1 GI:2402876
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 341)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mousees@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:556183
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 328.
Location/Qualifiers
1..341
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:975455"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse skin (#937313)"
/note="Organ: skin; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt, whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR vector; -5'
adaptor sequence: 5' GAATTCGGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTT 3'"

ORIGIN
Query Match 1.7%; Score 20; DB 9; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CTTCTCTCATCCGGGAGAGC 766
|||||
DB 141 CTTCTCTCATCCGGGAGAGC 160
|||||

RESULT 58
CB705122 347 bp mRNA linear EST 10-APR-2003
LOCUS AMGNNUC:NRHY3-00182-H5-A W Rat hypothalamus (10735) Rattus
DEFINITION norvegicus cDNA clone nrby3-00182-h5 5', mRNA sequence.
ACCESSION CB705122
VERSION CB705122.1 GI:29762270
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 347)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00182 row: h column: 5.
Location/Qualifiers
1..347
/organism="Rattus norvegicus"
/mol_type="mRNA"

```

/db_xref="taxon:10116"
/clone="nrhy3-00182-45"
/tissue_type="hypothalamus"
/clone_lib="W Rat hypothalamus (10735)"
/notes="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; W Rat
hypothalamus"

ORIGIN
Query Match      1.7%; Score 20; DB 14; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGGC 766
Db 15 CCTTCCTCATCCGGGAGGC 34

RESULT 59
AI585555
LOCUS
DEFINITION
vk63907.y1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
IMAGE:959388 5' similar to gb:M61906 PHOSPHATIDYLINOSITOL 3-KINASE
REGULATORY ALPHA SUBUNIT (HUMAN); gb:M60651 Mouse 85 kDa
phosphoprotein (MOUSE); mRNA sequence.
ACCESSION
AI585555
VERSION
AI585555.1 GI:4571452
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 376)
AUTHORS
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Peterson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,B., Kohn,S., Shin,T., Jackson,F., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
CONTACT: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:548180
This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation
Seq primer: -40RP from Gibco
High quality sequence stop: 355.
FEATURES
Location/Qualifiers
1..376
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:959388"
/tissue_type="Tcell"
/dev_stage="W30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse Tcell 937311"
/notes="Organ: Blood; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. M30 CD4+ cells. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGCAG
3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN
Query Match      1.7%; Score 20; DB 9; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 747 CCTTCCTCATCCGGGAGGC 766
Db 159 CCTTCCTCATCCGGGAGGC 178

RESULT 60
CB810150
LOCUS
DEFINITION
CB810150.1 GI:29931755
ACCESSION
CB810150
VERSION
CB810150.1
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 381)
AUTHORS
Angen EST Program
Angen Rat EST Program
Unpublished (2003)
COMMENT
Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00081 row: 9 column: 7.
FEATURES
Location/Qualifiers
1..381
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrpe4-00081-g7"
/tissue_type="placenta embryo"
/clone_lib="mrpe4 (10390)"
/notes="Vector: pSPORT1; Site_1: SalI; Site_2: NotI;
placenta embryo day 17"
ORIGIN
Query Match      1.7%; Score 20; DB 14; Length 381;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGGAGGC 766
Db 57 CCTTCCTCATCCGGGAGGC 76

RESULT 61
AA510624
LOCUS
DEFINITION
AA510624.r1 Soares mammary gland NbMVG Mus musculus cDNA clone
IMAGE:863515 5' similar to gb:M61906 PHOSPHATIDYLINOSITOL 3-KINASE
REGULATORY ALPHA SUBUNIT (HUMAN); gb:M60651 Mouse 85 kDa
phosphoprotein (MOUSE); mRNA sequence.
ACCESSION
AA510624
VERSION
AA510624.1 GI:2248478
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 388)
AUTHORS
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gaisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
CONTACT: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project

```


Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:507603
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 380.
Location/Qualifiers
1. .389
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:863515"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NbMMG"
/note="Organ: mammary gland; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo (dT) primer [5',
TGTTACCAATCTGAAGTCGGAGCGCCGCGAATGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."

ORIGIN
Query Match 1.7%; Score 20; DB 9; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CTTCTCATCCGGAGAGC 766
|||||
DB 158 CTTCTCATCCGGAGAGC 177
|||||

RESULT 62
CB802633
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CB802633 411 bp mRNA linear EST 16-MAY-2003
AMGNNUC:NRHW1-00297-F7-A W rat hypo+pit (10478) Rattus norvegicus
cDNA clone nrhw1-00297-f7 5', mRNA sequence.
CB802633
CB802633.1 GI:29917058
EST.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
1 (bases 1 to 411)
Amgen EST Program.
Amgen Rat EST Program
Unpublished (2003)
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00297 row: f column: 7.
Location/Qualifiers
1. .411
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrhw1-00297-f7"
/tissue_type="hypo-pit"

FEATURES
source

ORIGIN
Query Match 1.7%; Score 20; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CTTCTCATCCGGAGAGC 766
|||||
DB 63 CTTCTCATCCGGAGAGC 82
|||||

RESULT 63
AA271813
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA271813 425 bp mRNA linear EST 26-MAR-1997
va72h07.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone
IMAGE:736957 5', similar to gb:M61906 PHOSPHATIDYLINOSITOL 3-KINASE
REGULATORY ALPHA SUBUNIT (HUMAN); gb:M60651 Mouse 85 kDa
phosphoprotein (MOUSE); mRNA sequence.
AA271813
AA271813.1 GI:1910158
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 425)
Marra, M., Hallier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:454005
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 401.
Location/Qualifiers
1. .425
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:736957"
/sex="unknown"
/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
/clone_lib="Soares mouse 3NME12 5"
/note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was primed with a Not I - oligo (dT) primer
[5', TGTTACCAATCTGAAGTCGGAGCGCCGCGAATGTTTTTTTTTTTTTTTTTTT
3'], on total mouse RNA (provided by Minoru Ko, Wayne
State Univ.); double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 1.7%; Score 20; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

Qy 747 CCTCTCCTATCGGGAGAGC 766
|||||
Db 207 CCTCTCCTATCGGGAGAGC 226
|||||

RESULT 64
BF533481/c
LOCUS
DEFINITION
602074164F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4211198 5',
mRNA sequence.
ACCESSION BF533481
VERSION BF533481.1 GI:11620844
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
1 (bases 1 to 427)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue procurement: Jeffrey E. Green, M.D.
CDNA library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLN9779 row: P column: 15
High quality sequence stop: 426.
Location/Qualifiers
1..427
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4211198"
/clone_lib="NCI_CGAP_Li9"
/notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 1.7%; Score 20; DB 10; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 CTGGGCTTCCTCCCTGCG 75
|||||
Db 195 CTGGGCTTCCTCCCTGCG 176
|||||

RESULT 65
BE015226
LOCUS
DEFINITION
127451 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE015226
VERSION BE015226.1 GI:8276272
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
REFERENCE
1 (bases 1 to 439)
AUTHORS Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 747 CCTCTCCTATCGGGAGAGC 766
|||||
Db 207 CCTCTCCTATCGGGAGAGC 226
|||||

RESULT 64
BF533481/c
LOCUS
DEFINITION
602074164F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4211198 5',
mRNA sequence.
ACCESSION BF533481
VERSION BF533481.1 GI:11620844
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
1 (bases 1 to 427)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue procurement: Jeffrey E. Green, M.D.
CDNA library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLN9779 row: P column: 15
High quality sequence stop: 426.
Location/Qualifiers
1..427
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4211198"
/clone_lib="NCI_CGAP_Li9"
/notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 1.7%; Score 20; DB 10; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 CTGGGCTTCCTCCCTGCG 75
|||||
Db 195 CTGGGCTTCCTCCCTGCG 176
|||||

RESULT 65
BE015226
LOCUS
DEFINITION
127451 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE015226
VERSION BE015226.1 GI:8276272
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
REFERENCE
1 (bases 1 to 439)
AUTHORS Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,

```

```

Quackenbush,J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
PUBMED
12226715
COMMENT
Contact: Smith_TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 57 row: D column: 21
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1..439
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

ORIGIN
Query Match 1.7%; Score 20; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 377 TGAGTGTCTGTGCTGAGGAAC 396
|||||
Db 328 TGAGTGTCTGTGCTGAGGAAC 347
|||||

RESULT 66
CF391835
LOCUS
DEFINITION
RDR3_10_F11.b1.A022 Loblolly pine roots recovering from drought
D33 Pinus taeda cDNA clone RDR3_10_F11.A022 3', mRNA sequence.
ACCESSION CF391835
VERSION CF391835.1 GI:34350252
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
REFERENCE
1 (bases 1 to 449)
AUTHORS Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Dean,J.,
Gebremedhin,M., Dervinis,C., Martin,T., White,T., Davis,J. and
Neale,D.
TITLE An EST database from loblolly pine (Pinus taeda) roots recovering
from drought stress
JOURNAL Unpublished (2003)
COMMENT Other ESTs: RDR3_10_F11.g1.A022
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mprat@uga.edu
RNA prepared and library constructed by W. Walter Lorenz, School of
Forestry, University of Georgia; plant material prepared at the
University of Florida; sequencing done in the laboratory for
Genomics and Bioinformatics, University of Georgia. Sequence ends

```

have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
Seq primer: M13-21 (TGTAACACGCGCCAGT)
POLYA=yes.

FEATURES

Source

Location/Qualifiers

```
1..449
/organism="Pinus taeda"
/mol_type="mRNA"
/strain="CCLONES"
/db_xref="taxon:3352"
/clone="RTDR3_10_F11_A022"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Loblolly pine roots recovering from drought
DR3"
/note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The
library was prepared from polyA+ RNA from loblolly pine
(Pinus taeda) roots recovering from drought. Water was
withheld from ramet clones until predawn needle water
potential reached -1.75 MPa. Plants were well watered on
day 7 and allowed to recover for 2 days, at which time
roots were harvested for RNA isolation. Double-stranded
cDNA was cloned unidirectionally into pSL1180. Inserts
excised with EcoRI (5' end) and XhoI (3' end)."
```

ORIGIN

```
Query Match 1.7%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 412 CAGCAGAGAAAATCTCTGC 431

DB 192 CAGCAGAGAAAATCTCTGC 211

RESULT 67

```
BQ198515 455 bp mRNA linear EST 07-MAY-2003
LOCUS NXLV132_C04_F NXLV (Nsf Xylem Late wood vertical) Pinus taeda cDNA
DEFINITION clone NXLV132_C04_5', mRNA sequence.
```

ACCESSION BQ198515.1 GI:20384477

VERSION BQ198515

KEYWORDS EST.

SOURCE Pinus taeda (loblolly pine)

ORGANISM

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
```

REFERENCE Sederoff, R.

TITLE Molecular Basis of Wood Formation in the Pine Megagenome

JOURNAL Unpublished (2000)

COMMENT Contact: Sederoff, Ron

Forest Biotechnology

North Carolina State University

840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,

NC 27695, USA

Tel: 919 515 7800

Fax: 919 515 7801

Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu

Please see <http://web.anc.umn.edu/biodata/nsfpine/> for further information.

Seq primer: T3.

Location/Qualifiers

FEATURES

Source

```
1..455
/organism="Pinus taeda"
/mol_type="mRNA"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXLV132_C04"
/tissue_type="primary xylem"
/dev_stage="late wood"
/lab_host="X11-Blue"
/clone_lib="NXLV (Nsf Xylem Late wood Vertical)"
```

/note="Vector: pTriplex; Site 1: EcoRI; The library is from late (summer-August) wood, taken from below the crown of a 20 year old tree. The harvested xylem tissue was on the cusp between transitional and mature wood. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATCGCCATTATGGCC'."

ORIGIN

```
Query Match 1.7%; Score 20; DB 13; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 412 CAGCAGAGAAAATCTCTGC 431

DB 169 CAGCAGAGAAAATCTCTGC 188

RESULT 68

```
AW783749 462 bp mRNA linear EST 10-MAY-2001
LOCUS rp0se02.y1 Kloeck Brook Zeldia punctata SL1 Zeldia punctata cDNA 5',
DEFINITION similar to WP:D1007.6 CE09041 40S RIBOSOMAL PROTEIN S10 ;, mRNA
sequence.
```

ACCESSION AW783749.1 GI:7798346

VERSION AW783749

KEYWORDS EST

SOURCE Zeldia punctata

ORGANISM

```
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Cephaloboidae; Cephalobidae; Zeldia.
```

REFERENCE 1 (bases 1 to 462)

AUTHORS

McCarter, J., Clifton, S., Chiapelli, B., Page, D., Martin, J.,
Wyllie, T., Dente, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, F.,
Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCaun, R., Waterston, R. and Wilson, R.
The Washington Univ. Nematode EST Project, 1999

TITLE

JOURNAL

COMMENT

Contact: McCarter, JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Merry Brook and Dr. Andrew Kloeck
(kloeck@biology.wustl.edu) at Divergence LLC. DNA sequencing by:
Washington University Genome Sequencing Center St. Louis.

Seq primer: SL1 primer

High quality sequence stop: 449.

Location/Qualifiers

FEATURES

source

1..462

/organism="Zeldia punctata"

/mol_type="mRNA"

/strain="PDL3"

/db_xref="taxon:49351"

/dev_stage="Mixed"

/lab_host="Top10P"

/clone_lib="Kloeck Brook Zeldia punctata SL1"

/note="Vector: pCRII-TOPO; Site 1: NA; Site 2: NA;
SL1-Oligo(dT) PCR-based library. Zeldia cDNA PCR products
containing SL1 on the 5' end and oligo(dT) on the 3' end
were non-directionally cloned into pCRII-TOPO (Invitrogen)
following the TOPO TA cloning protocol. The cDNA insert
can be excised by digestion with EcoRI. This library was
constructed by Merry Brook and Dr. Andrew Kloeck at
Divergence LLC. The strain PDL3 used in this work was
provided by the Caenorhabditis Genetics Center at the
University of Minnesota."

ORIGIN


```

1. 477
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2989178"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NHG MGC 8"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

source
 Query Match 1.7%; Score 20; DB 10; Length 477;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
 QY 747 CTTCTCTCATCGGAGAGC 766
 Db 157 CTTCTCTCATCGGAGAGC 176
 RESULT 73
 CA545265 485 bp mRNA linear EST 19-NOV-2002
 LOCUS K0110G10.5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit+/Sca-1-)
 DEFINITION cDNA Library (Long) Mus musculus cDNA clone NIA.K0110G10
 IMAGE:30040401 5', mRNA sequence.
 CA545265
 CA545265.1 GI:25088056
 EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 485)
 Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
 Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H.
 Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
 (Lin-/c-Kit+/Sca-1-) cDNA Library (Long)
 Unpublished (2001)
 JOURNAL Other ESTs: K0110G10-3
 COMMENT Laboratory of Genetics
 Contact: Dawood B. Dudekula
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: K0110 row: G column: 10
 Seq primer: M13 Reverse
 High quality sequence stop: 485
 POLYA=No.

FEATURES
 Location/Qualifiers
 1..485
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6Ncr"
 /db_xref="taxon:10090"
 /clone="NIA:K0110G10 IMAGE:30040401"
 /tissue_type="Hematopoietic Stem Cell"
 (Lin-/c-Kit+/Sca-1-)
 /dev_stage="Age approx. 10 weeks old"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Hematopoietic Stem Cell"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
 obtained from Drs. Dennis Taub, Dan Longo (National
 Institute on Aging, USA), Jonathan Keller (National Cancer
 Institute, USA). Double-stranded cDNAs were synthesized
 with an Oligo(dT) primer (Invitrogen).
 5'-PGACTAGTTCAGATCCGACGCGCCCTTTT-3' from
 2.4 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Lona-linker Lh-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The

```

1. 478
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:959388"
/tissue_type="Tcell"
/dev_stage="W30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse Tcell 937311"
/note="Organ: blood; Vector: pBluescript SK-; Site 1:  
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. M30 CD4+ cells. Average insert size: 1.0 Kb;
```

source
 Query Match 1.7%; Score 20; DB 10; Length 477;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
 QY 747 CTTCTCTCATCGGAGAGC 766
 Db 248 CTTCTCTCATCGGAGAGC 267
 RESULT 72
 AA547055 478 bp mRNA linear EST 05-AUG-1997
 LOCUS vk63g07.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
 DEFINITION IMAGE:959388 5', similar to gb:M61906 PHOSPHATIDYLINOSITOL 3-KINASE
 REGULATORY ALPHA SUBUNIT (HUMAN); gb:M60651 Mouse 85 kDa
 phosphoprotein (MOUSE);, mRNA sequence.
 AA547055
 AA547055.1 GI:2308346
 EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 478)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 JOURNAL Contact: Maria M/Mouse EST Project
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:548180
 High quality sequence stop: 145.

products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The D10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 Kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 1.7%; Score 20; DB 14; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 GTGGCCCTGGCGAGTTTCCC 528
|||||
Db 426 GTGGCCCTGGCGAGTTTCCC 445

RESULT 74

AA051252

LOCUS

DEFINITION mJ43c04.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
clone IMAGE:478854 5', mRNA sequence.

ACCESSION

AA051252

VERSION

AA051252.1

KEYWORDS

EST.

SOURCE

ORGANISM

Mus musculus

(house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 507)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisler, S., Kucaba, J., Lucy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

MGI:289598

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 483.

Location/Qualifiers

1, 507

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:478854"

/sex="unknown"

/tissue type="embryo"

/dev stage="13.5-14.5dpc total fetus"

/lab_host="DH10B"

/clone_lib="Soares mouse embryo NbME13.5 14.5"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5',

TGTTACCAATCTGAAGTGGAGCGCGCGGAATTTTTTTTTTTTTTT

T 3'], on equal amounts of mRNA from 2 13.5dpc and 2

14.5dpc embryos [total RNA provided by Minoru KO, Wayne

State Univ., from 2]; double-stranded cDNA was ligated to

Eco RI adaptors (Pharmacia), digested with Not I and

cloned into the Not I and Eco RI sites of the modified

pT73 vector. Library went through one round of

normalization, and was constructed by Bento Soares and

ORIGIN

Query Match 1.7%; Score 20; DB 9; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 GTGGCCCTGGCGAGTTTCCC 528
|||||
Db 449 GTGGCCCTGGCGAGTTTCCC 468

RESULT 75

AUI27488

LOCUS

DEFINITION AUI27488 NT2RP2 Homo sapiens cDNA clone NT2RP2001397 5', mRNA
sequence.

ACCESSION

AUI27488

VERSION

AUI27488.1

KEYWORDS

EST.

SOURCE

Homo sapiens

(human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 525)

Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,

Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and

Isogai, T.

HRI human cDNA project

Unpublished (2000)

Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

Location/Qualifiers

1, 525

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RP2001397"

/cell_type="teratocarcinoma"

/cell_line="NT2"

/clone_lib="NT2RP2"

/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor

cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Query Match 1.7%; Score 20; DB 9; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 GAAGAAAATCTCTGCCAAGC 436
|||||
Db 452 GAAGAAAATCTCTGCCAAGC 471

RESULT 76

AA818383/c

LOCUS

DEFINITION AA818383 526 bp mRNA linear EST 03-JUL-1999
UI-R-A0-at-g-02-0-UI.s1 UI-R-A0 Rattus norvegicus cDNA clone
UI-R-A0-at-g-02-0-UI 3', similar to dbj|D64046|RATP13KB Rat mRNA for
phosphatidylinositol 3-kinase p85 beta subunit, complete cds, mRNA
sequence.

ACCESSION

AA818383

VERSION

AA818383.1

KEYWORDS

EST.

SOURCE ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 526)
REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.
AUTHORS Normalization and subtraction: two approaches to facilitate gene
TITLE discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PubMed 8889548
COMMENT On Feb 17, 1998 this sequence version replaced gi:2889122.
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence tag present in the cDNA between the NotI site and the
oligo-dT track served to identify it as a clone from the normalized
adult Brain library. cDNA Library Preparation: M. Fatima Bonaldo,
Ph.D. Clone distribution: clones will be available through Research
Genetics This clone is also available through the I.M.A.G.E.
Consortium at LLNL (info@image.llnl.gov). IMAGE ID=1768058
Seq primer: M13 Forward
POLYA=No.
FEATURES source
1..526
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-AO-at-g-02-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-AO"
/note="Vector: pTY73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; This library
consists of a mixture of individually tagged normalized
libraries constructed from rat placenta, adult lung,
brain, liver, kidney, heart, spleen, ovary, and muscle.
The tag is a string of 3-5 nucleotides present between the
Not I site and the oligo-dT track which allows
identification of the library of origin of a clone within
the mixture."
ORIGIN
Query Match 1.7%; Score 20; DB 9; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.9e-02; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;
QY 747 CTTCTCTCATCGGGAGAGC 766
Db 312 CTTCTCTCATCGGGAGAGC 293
|||||
CD892975 549 bp mRNA linear EST 14-JUL-2003
G118_122104F010726 G118 Triticum aestivum cDNA clone G118122104,
mRNA sequence.
ACCESSION CD892975
VERSION CD892975.1 GI:32663549
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 549)
REFERENCE
AUTHORS Genoplante.

TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).
FEATURES source
1..549
Location/Qualifiers
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G118122104"
/tissue_type="grain (118 degrees per day after
pollination)"
/clone_lib="G118"
ORIGIN
Query Match 1.7%; Score 20; DB 14; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.9e+02; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;
QY 32 CTCTGTGACAGAGCTCAAG 51
Db 371 CTCTGTGACAGAGCTCAAG 390
|||||
RESULT 78
AW647304 552 bp mRNA linear EST 04-APR-2000
LOCUS EST325847 R3DA Mus musculus cDNA clone R3DA172, mRNA sequence.
DEFINITION AW647304
ACCESSION AW647304
VERSION AW647304.1 GI:7408532
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 552)
AUTHORS Earle-Hughes, J., Cho, J., Hansen, T.S., Lee, H.H., Quackenbush, J.,
Adams, M.D., Fraser, C.M. and Venter, J.C.
Generation of ESTs from Murine adipose, differentiated 373 cell
line
JOURNAL Unpublished (1994)
COMMENT Other ESTs: TC68962
Contact: Julie Earle-Hughes
TIGR
The Institute For Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208.
FEATURES source
1..552
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/cultivar="3T3"
/db_xref="taxon:10090"
/clone="R3DA172"
/tissue_type="adipose"
/clone_lib="R3DA"
/note="Site_1: EcoRI; Site_2: XhoI"
ORIGIN
Query Match 1.7%; Score 20; DB 10; Length 552;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1076 TCTGAGCTGCCACAGGGA 1095
|||||

```

Db      503 TCTGAAGCTGCCACAGGGA 522

RESULT 79
BJ632593      553 bp mRNA linear EST 01-OCT-2003
LOCUS      BJ632593 NIBB Mochii normalized Xenopus early gastrula library
DEFINITION  Xenopus laevis cDNA clone XL175el9 3', mRNA sequence.
ACCESSION   BJ632593
VERSION     BJ632593.1 GI:372833952
KEYWORDS
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM    Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1 (bases 1 to 553)
AUTHORS    Kitayama,A., Terabaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
TITLE      Expressed genes in X. laevis embryo
JOURNAL    Unpublished (2001)
COMMENT    Contact: Tadasi Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            The information of this clone is available through the following
            URL:
            http://xenopus.nibb.ac.jp.

FEATURES             Location/Qualifiers
     source           1..553
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8335"
                     /clone="XL175el9"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 10.5"
                     /clone_lib="NIBB Mochii normalized Xenopus early gastrula
                     library"

ORIGIN
Query Match      1.7%; Score 20; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      825 ACCGGATCAGACTACAGG 844
Db      496 ACCGGATCAGACTACAGG 515

RESULT 80
CB586549      582 bp mRNA linear EST 03-APR-2003
LOCUS      CB586549 AMGNNUC-NRHY4-00132-H12-A W Rat hypothalamus (10464) Rattus
DEFINITION  norvegicus cDNA clone nrhy4-00132-h12 5', mRNA sequence.
ACCESSION   CB586549
VERSION     CB586549.1 GI:29530590
KEYWORDS
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 582)
AUTHORS    Amgen EST Program.
TITLE      Amgen Rat EST Program
JOURNAL    Unpublished (2003)
COMMENT    Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881

FEATURES             Location/Qualifiers
     source           1..582
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L10choCK0-4-B10"
                     /sex="M"
                     /cell_line="Cho-CK"
                     /lab_host="Tcpl0P"
                     /clone_lib="L10choCK0"
                     /notes="Organ: Liver; Vector: pT7T3-Pac; Site:1: EcoRI;
                     Site 2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996). Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."

ORIGIN
Query Match      1.7%; Score 20; DB 14; Length 602;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      747 CCTTCCTCATCGGGAGAGC 766
Db      747 CCTTCCTCATCGGGAGAGC 766

```


Plate: H4019 row: E column: 02
Seq primer: -21M13 Reverse
High quality sequence stop: 640
POLYA=No.

FEATURES
source
1. 640
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="niaest:H4019E02-5"
/db_xref="taxon:10090"
/clone="H4019E02"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 7,407 clones from more than 20 cDNA libraries."

ORIGIN

Query Match 1.7%; Score 20; DB 13; Length 640;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 GTGGCCCTGGCAGTTTCCC 528
Db 426 GTGGCCCTGGCAGTTTCCC 445

RESULT 85

BY746881

LOCUS

DEFINITION BY746881 RIKEN full-length enriched, 2 days neonate thymus thymic

cells (NOD) Mus musculus cDNA clone E430023D24 5', mRNA sequence.

BY746881.1 GI:27174830

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 653)

Okazaki, Y., Furum, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bratt, D., Brusic, V.,

Chochia, C., Corbani, D.E., Cousins, S., Dalla, E., Dragani, T.A.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,

Kurochkin, I. V., Lee, Y., Leinhardt, B., Lyons, P. A., Maglott, D. R.,

Maltais, L., Marchionni, B., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J. C., Reid, D. J., Reid, J., Ring, B. Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tonita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,

Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

MEDLINE

PUBMED

COMMENT

22354683
12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.c.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,

Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,

Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,

Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,

Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,

Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics

Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome

Trust/MRC building Addenbrookes Hospital Cambridge) whose

assistance we gratefully acknowledge.

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

Location/Qualifiers

1. 653

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NOD"

/db_xref="taxon:10090"

/clone="E430023D24"

/cissue_type="thymus"

/cell_type="thymic cells"

/clone_lib="RIKEN full-length enriched, 2 days neonate

thymus thymic cells (NOD)"

ORIGIN

Query Match 1.7%; Score 20; DB 13; Length 653;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 GTGGCCCTGGCAGTTTCCC 528

Db 466 GTGGCCCTGGCAGTTTCCC 485

RESULT 86

BY746881

LOCUS

DEFINITION

BY746881 RIKEN full-length enriched, 2 days neonate thymus thymic

cells (NOD) Mus musculus cDNA clone E430023D24 5', mRNA sequence.

BY746881.1 GI:24992553

ACCESSION

VERSION

KEYWORDS

SOURCE

Clona intestinalis

Clona intestinalis

Eukaryota; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.

1 (bases 1 to 654)
 Satochi, Shin-I., Kohara, Y. and Satochi, N.
 Expressed genes in Ciona intestinalis (2002c)
 Unpublished (2002)
 Contact: Nori Satochi
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satochi@ascidian.zool.kyoto-u.ac.jp.

Location/Qualifiers

1..654
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cint021h18"
 /tissue_type="heart"
 /clone_lib="Nori Satochi unpublished cDNA library, heart"

ORIGIN

Query Match 1.7%; Score 20; DB 13; Length 654;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 847 CCATGCTTGCATGCT 866
 |||||
 DB 381 CCATGCTTGCATGCT 362

RESULT 87
 BE261614
 LOCUS
 DEFINITION 601149190F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3501670 5', mRNA linear EST 26-CCT-2000
 mRNA sequence.

ACCESSION BE261614
 VERSION BE261614.1 GI:9133326
 KEYWORDS EST.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 655)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCW172 row: d column: 23
 High quality sequence stop: 637.
 Location/Qualifiers

1..655
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3501670"
 /tissue_type="neuroblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_19"
 /note="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.7%; Score 20; DB 10; Length 655;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCGGAGAGC 766
 |||||
 DB 628 CCTTCCTCATCGGAGAGC 647

RESULT 88
 AI510095
 LOCUS
 DEFINITION m43C04.Y1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:478854 5', similar to TR:Q13239 Q13239 PUTATIVE SRC-LIKE ADAPTER PROTEIN ;, mRNA sequence.

ACCESSION AI510095
 VERSION AI510095.1 GI:4409000
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 670)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 This read is a RESEQUENCE of a previously sequenced mouse clone. This read has been verified (found to hit its original self in the correct orientation)
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gibco
 High quality sequence stop: 460
 POLYA=No.

Location/Qualifiers

1..670
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:478854"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTCAAGTGGGCGGCGGGAATTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and

```

ORIGIN
M.Fatima Bonaldo. "
Query Match 1.7%; Score 20; DB 9; Length 670;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 GTGGCCCTGGGAGTTTCCC 528
Db 449 GTGGCCCTGGGAGTTTCCC 468

RESULT 89
BW305486/c
LOCUS BW305486 689 bp mRNA linear EST 11-NOV-2002
DEFINITION BW305486 Nori Satoh unpublished cDNA library, heart Ciona
intestinalis cDNA clone c1nt010j11 5', mRNA sequence.
ACCESSION BW305486
VERSION BW305486.1 GI:24886097
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 689)
AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..689
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="c1nt010j11"
/tissue_type="heart"
/clone_lib="Nori Satoh unpublished cDNA library, heart"

ORIGIN
Query Match 1.7%; Score 20; DB 13; Length 689;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 847 CCACCTGCCTTGACATGGCT 866
Db 288 CCACCTGCCTTGACATGGCT 269

RESULT 90
BZ828172/c
LOCUS BZ828172 696 bp DNA linear GSS 18-MAR-2003
DEFINITION PUFGR687D ZM 0.6-1.0-KB Zea mays genomic clone ZM07A315K16,
genomic survey sequence.
ACCESSION BZ828172
VERSION BZ828172.1 GI:29049444
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 696)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Bennett,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)

COMMENT
Other GSSs: PUFGR687B
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..696
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM07A315K16"
/clone_lib="ZM_0.6-1.0_KB"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN
Query Match 1.7%; Score 20; DB 28; Length 696;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TCCCTCCCTGGCTCGGCTGT 83
Db 594 TCCCTCCCTGGCTCGGCTGT 575

RESULT 91
AV968086/c
LOCUS AV968086 710 bp mRNA linear EST 14-MAR-2002
DEFINITION AV968086 Nori Satoh unpublished cDNA library, larva Ciona
intestinalis cDNA clone c1vl4h05 5', mRNA sequence.
ACCESSION AV968086
VERSION AV968086.1 GI:19457850
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 710)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..710
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="c1vl4h05"
/tissue_type="whole animal"
/dev_stage="larva"
/clone_lib="Nori Satoh unpublished cDNA library, larva"

ORIGIN
Query Match 1.7%; Score 20; DB 9; Length 710;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 847 CCACCTGCCTTGACATGGCT 866
Db 103 CCACCTGCCTTGACATGGCT 84

```

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFZp686B23241) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
Location/Qualifiers
1..719
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686B23241"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hicc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match 1.7%; Score 20; DB 13; Length 719;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 472 TGTGACCATGGAAGCAGAGA 491
Db 631 TGTGACCATGGAAGCAGAGA 612
|||||
RESULT 94
BI153749
LOCUS
DEFINITION
602871407F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5003328 5';
mRNA sequence.
BI153749
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11039 row: n column: 01
High quality sequence stop: 696.
Location/Qualifiers
1..736
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5003328"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam2"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sall;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFZp686B23241) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
Location/Qualifiers
1..710
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4341278"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: Sall; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 1.7%; Score 20; DB 10; Length 710;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 747 CTTCTCTATCGGAGAGC 766
Db 557 CTTCTCTATCGGAGAGC 576
|||||
RESULT 93
BX642175/c
LOCUS
DEFINITION
BX642175.1 GI:12102078
EST. EST 04-SEP-2003
DKFZp686B23241 r1 686 (synonym: hicc3) Homo sapiens cDNA clone
DKFZp686B23241 5', mRNA sequence.
BX642175
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Wiemann,S.
EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)
Unpublished (2003)
Contact: MIFS
MIFS

/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.7%; Score 20; DB 10; Length 756;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 TCTTCATCCGGAGAGCCAG 769
DB 650 TCTTCATCCGGAGAGCCAG 631

RESULT 98

BG873348
LOCUS 602794320F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4925596 5',
DEFINITION mRNA sequence.

ACCESSION BG873348
VERSION BG873348.1 GI:14223888
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 763)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10850 row: g column: 05
High quality sequence stop: 729.

FEATURES

source
1..763
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="FVB/N"
/clone="IMAGE:4925596"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 1.7%; Score 20; DB 12; Length 763;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGAGAGC 766
DB 55 CCTTCCTCATCCGGAGAGC 74

RESULT 99

BI084247
LOCUS 764 bp mRNA linear EST 20-JUN-2001

DEFINITION

602859751F1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5014285 5',
mRNA sequence.

ACCESSION BI084247
VERSION BI084247.1 GI:14502577
KEYWORDS EST.

SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 764)

REFERENCE

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLAM1821 row: f column: 14
High quality sequence stop: 762.

FEATURES

source
1..764
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5014285"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"

/note="Organ: salivary gland; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.7%; Score 20; DB 12; Length 764;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCCTCATCCGGAGAGC 766
DB 745 CCTTCCTCATCCGGAGAGC 764

RESULT 100

BZ774469
LOCUS BZ774469 773 bp DNA linear GSS 13-MAR-2003

DEFINITION mcv95e04.g1 HFOSMID007 Homo sapiens genomic, genomic survey sequence.

ACCESSION BZ774469
VERSION BZ774469.1 GI:28948153
KEYWORDS GSS.

SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 773)

REFERENCE

AUTHORS Cook, L., Delehaunty, K., Fewell, G., Fulton, L., Magrini, V., Mardis, E., Miner, T., Nash, W., Williams, D. and Wilson, R.K.
TITLE Homo sapiens Fosmid End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson

Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu

Plate: mcv95 row: e column: 04
 Class: fosmid ends
 High quality sequence start: 36
 High quality sequence stop: 550.

FEATURES

source

1..773
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone_lib="HFOSMID007"
 /notes="Vector: pcc01fos; Site 1: Eco72i; Human whole
 genome fosmid library was prepared at Washington
 University Genome Sequencing Center. DNA was sheared for
 blunt-ended ligation into pcc01fos inducible vector. DNA
 was ordered from Coriell Cell Repository's DNA
 polymorphism discovery resource."

ORIGIN

Query Match 1..7% Score 20; DB 28; Length 773;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GACAATGGCTGGCTGTACAT 876
 |||||
 Db 289 GACAATGGCTGGCTGTACAT 308

Search completed: March 25, 2004, 08:38:57
 Job time : 3522 secs